

AD _____

Award Number: DAMD17-99-1-9154

TITLE: Model and Expansion Based Methods of Detection of Small Masses in Radiographs of Dense Breasts

PRINCIPAL INVESTIGATOR: Andrew F. Laine, D.Sc.

CONTRACTING ORGANIZATION: Columbia University
New York, NY 10027

REPORT DATE: December 2003

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20040820 025

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE December 2003	3. REPORT TYPE AND DATES COVERED Final (1 Jun 1999 - 30 Nov 2003)	
4. TITLE AND SUBTITLE Model and Expansion Based Methods of Detection of Small Masses in Radiographs of Dense Breasts		5. FUNDING NUMBERS DAMD17-99-1-9154	
6. AUTHOR(S) Andrew F. Laine, D.Sc.			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Columbia University New York, NY 10027 E-Mail: LAINE@columbia.edu		8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012		10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES Original contains color plates: ALL DTIC reproductions will be in black and white			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited		12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 Words) Our goal is to identify representations to assist in the detection masses in dense mammograms having a diameter less than 1 cm. The central idea of this project is to detect subtle masses by tuning the central frequency and width of a basis function used in an overcomplete expansion. By modeling the shape of a mass through this flexibility we hope to detect small and subtle masses in dense breasts and improve the chances of early detection in screening mammography. During this final year of the project we implemented a level-set method of segmentation that made use of a local homogeneity operator for the detection of subtle masses in digital mammography. These methods are currently integrated into our previously described multi-scale expansion framework and will be tested using an existing public database of mammograms with ground truth of known disease.			
14. SUBJECT TERMS Detection, Masses, Multiscale Analysis, Wavelets, CAD		15. NUMBER OF PAGES 46	
		16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Table of Contents

Cover.....	1
SF 298.....	2
Table of Contents.....	3
Introduction.....	4
Body.....	4
Key Research Accomplishments.....	11
Reportable Outcomes.....	11
Conclusions.....	13
References.....	14
Appendices.....	15

1. INTRODUCTION

In the first year of this study we showed that the dyadic transform was not sufficient for our case mass of detection [1]. Last year, we developed multi-scale adaptive histogram equalization (MSAHE) that achieved a global contrast enhancement by adjusting contrast locally through reassigning a central pixel the value through a local histogram equalization mapping function. This past six months we have implemented a level set method of segmentation for the detection of masses within a multi-scale expansion. This report provides an overview of this method and describes progress to date. Since this is a final report, we will summarize previous accomplishments and status of the project overall, at the end.

Traditional methods of segmentation, such as pixel-based clustering, region growing, and edge detection, requires additional pre-processing and post-processing as well as a considerable amounts of expert intervention or information of the objects of interest. Furthermore the subsequent analysis of segmented objects is hampered by primitive, pixel or voxel level representations from region-based segmentation [1].

Deformable models, on the other hand, provide an explicit representation of the boundary and the shape of an object. They combine several desirable features such as inherent connectivity and smoothness, which counteract noise and boundary irregularities, as well as the ability to incorporate knowledge about the object of interest [1,3] [4]. However, parametric deformable models have two main limitations. First, in situations where the initial model and desired object boundary differ greatly in size and shape, the model must be re-parameterized dynamically to faithfully recover the object boundary. The second limitation is that it has difficulty dealing with topological adaptation such as splitting or merging model parts, a useful property for recovering either multiple objects or an object with unknown topology. This difficulty is caused by the fact that a new parameterization must be constructed whenever the topology change occurs, which requires sophisticated schemes [5, 6].

In the body of this report an alternative approach of coping with weaknesses in existing methods of segmentation of masses is described.

2. BODY

DETECION OF MASSES VIA EVOLUTION OF LEVEL SET BOUDARIES

In existing level-set methods, the gradient information is used as a stopping criterion for curve evolution, and also provides the attracting force to the zero level-set from the target boundary. However, in a discrete implementation, the gradient-based term can never fully stop the level-set evolution even for ideal edges, leakage is often unavoidable. Also the effective distance of the attracting force and blurring of edges become a trade-off in choosing the shape and support of the smoothing filter. The proposed homogeneity measurement provides easier and more robust edge estimation, and the possibility of fully stopping the level-set evolution. The homogeneity term decreases from a homogenous region to the boundary, which dramatically increases the effective

distance of the attracting force and also provides an additional measurement of the overall approximation to the target mass boundary. Therefore, it provides a reliable criterion of adaptively changing the advent speed. By using this term, the leakage problem was avoided effectively in most cases compared to traditional level-set methods. The computation of the homogeneity operator is fast and can be done within seconds on a PC workstation.

Curvature evolution and Level set

Level set segmentation [7, 8], also referred as geometric deformable models, provides an elegant solution to address the primary limitations of parametric deformable models. In the 2D case, the boundary of an object is implicitly represented as the zero-level set of a time dependent 2D function, which is usually called the level set function. A useful property of this approach is that the level set function remains a valid function while the embedded curve (the zero level set) can change its topology.

The evolution equation for the level set function $\phi(x, y, t)$ takes on the following formula [9]:

$$\frac{\partial \phi}{\partial t} + F |\nabla \phi| = 0 \quad (1)$$

The evolution of the level set function was determined by the speed function F . As an example, imagine that given an initial closed curve that is evolving under three simultaneous motions. First, it is expanding with a constant speed in its normal direction; second, it is moving with a speed proportional to its curvature; third, it is being passively advected by an underlying velocity field whose direction and strength depend on position and time, but not on the front itself. This entire motion can then be written in terms of the speed function as an explicit level set scheme:

$$F = F_{prop} + F_{curv} + F_{adv} \quad (2)$$

where $F_{prop} = F_0$ is the propagation expansion speed, $F_{curv} = -\varepsilon\kappa$ is the dependence of the speed on the curvature κ , and $F_{adv} = \vec{U}(x, y, t) \cdot \vec{n}$ is the advection speed, where \vec{n} is the normal to the front. The PDE in (1) can be solved with entropy-satisfying schemes given the speed function.

A small modification version of (2) gives the general formula for level set segmentation:

$$\phi_t + g_I(1 - \varepsilon\kappa) |\nabla \phi| - \beta \nabla P \cdot \nabla \phi = 0 \quad (3)$$

In the term $1 - \varepsilon\kappa$, the uniform expansion with speed 1 corresponds to the inflation force used by Cohen [10]. The diffusive term $\varepsilon\kappa$ smoothes out the high curvature regions and has the same regularizing effect as the internal deformation energy term in thin-plate-membrane splines [2]. The term g_I was computed from the image data, and provide a halting criterion for the speed function, the value of g_I should be between 0 and 1, and ideally, with 0 on the boundary and 1 within the homogeneous region (either within or outside the object). Typically, it can be estimated by the gradient:

$$g_I(x, y) = \frac{1}{1 + |\nabla(G_\sigma * I(x, y))|} \quad (4)$$

where the expression $G_\sigma * I$ denotes the image convolved with a Gaussian smoothing filter whose characteristic width is σ . The term $\nabla(G_\sigma * I(x, y))$ is essentially zero except where the image gradient changes rapidly, in which case the value becomes large. Thus, g_I is close to unity away from

boundaries and drops to zero near sharp changes in the image gradient. These changes presumably correspond to the edges of the desired shape. In other words, the first term in (3) anticipates steep drops in the image gradient, and retards the evolving front from passing out of the desired region. The second term in (3) is a force which attracts the surface towards the boundary, which has a stabilizing effect, especially when there is a large variation in the image gradient value. This term denotes the projection of an attractive force vector on the surface normal. This force, introduced in [11], is realized as the gradient of a potential field. Here:

$$P = -|\nabla(G_\sigma * I)| \quad (5)$$

attracts the surface to the edges in the image, the coefficient β controls the strength of this attraction.

Apparently, both terms depend on the edge map $\nabla(G_\sigma * I)$, and the quality of this edge estimation determines performance of the segmentation.

“Scale Map” based on homogeneity measurement

“Scale” is a fundamental, well-established concept in image processing [12, 13]. The premise behind this concept is to consider the local size of the object in carrying out whatever local operations that is to be carried out on the image. It has previously been used as a metric of local homogeneity [13].

“Object Scale” in an image C at any pixel c was defined as the radius $r(c)$ of the largest hyper ball centered at c which lies entirely in the object region [14].

A hyper ball $B_r(c)$ centered at pixel c is a collection of pixels around c , i.e. $B_r(c) = \{e \in C \mid \|c - e\| \leq r\}$.

For a ball $B_k(c)$ of any radius k centered at c , a fraction function $FO_k(c)$ was defined, which indicates whether the fraction of the ball boundary is sufficiently homogeneous to the inside region of the ball:

$$FO_k(c) = \frac{\sum_{d \in B_k(c) - B_{k-1}(c)} W(|f(c) - f(d)|)}{|B_k(c) - B_{k-1}(c)|}, \quad (6)$$

where $|B_k(c) - B_{k-1}(c)|$ is the number of pixels in $B_k(c) - B_{k-1}(c)$ and W is the homogeneity function, which measures the similarity of two pixels based on their pixel value in the image $f(x)$. Some typically used homogeneity functions for this sake are illustrated in Figure 1.

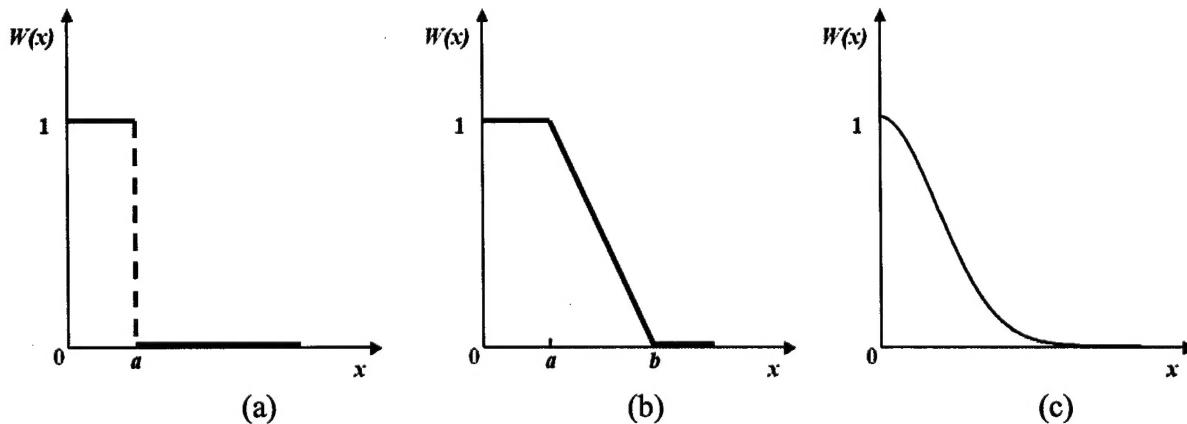


Figure 1: Examples of homogeneity functions for computing the fraction function in (6).

The mathematical formulas for the homogeneity functions in Figure 1 are:

$$(a) W(x) = \begin{cases} 1, & 0 \leq x \leq a \\ 0, & x > a \end{cases}$$

$$(b) W(x) = \begin{cases} 1, & 0 \leq x \leq a \\ \frac{b-x}{b-a}, & a \leq x \leq b \\ 0, & x > b \end{cases} \quad (7)$$

$$(c) W(x) = e^{-x^2/2k^2}, k > 0$$

By setting a fixed threshold $0 < t < 1$, the scale map at a pixel c can be computed by the following pseudo-code:

```

begin
  k=1;
  while FOk(c) ≥ t do
    k=k+1;
  endwhile
  r(c)=k;
end

```

Methodology for computing an edge map based on homogeneity metrics

The scale map described above provides a robust homogeneity measurement by incorporating a tolerance level t . For example, if we use $t = 7/8$, in a 3×3 neighborhood of a pixel c in a 2D scene, we allow one out of the eight neighboring pixels to belong to a different object (to account for noise) but still consider the neighborhood to be entirely within the same object. This actually provides a mechanism of denoising within the homogeneity measurement.

An edge was defined as a region of an image in which the pixel value changes significantly over a short distance [15]. Therefore the edge represents a region that has significant lower homogeneity. Thus a homogeneity measurement certainly provides edge information within an image.

Figure 2 shows the “scale map” computed by the algorithm above using different parameters.

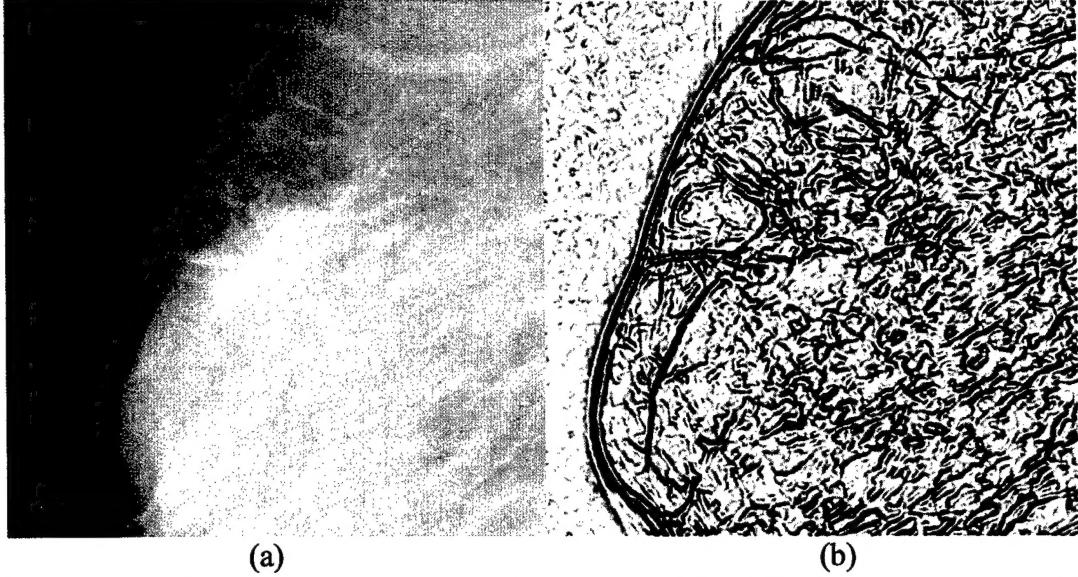


Figure 2: scale map of breast radiograph image. (a) original mammogram (256 gray levels). (b) scale map computed with parameter, $k^2 = 500, t=0.8$.

Here we use equation 7(c) as the homogeneity function for computation. The two parameters for computing the scale map are the shape parameter of the homogeneity function (k), and the tolerance value for homogeneity measurement (t). Parameter k determines how much variation in the pixel value is tolerated in terms of homogeneity. Parameter t , as discussed previously, determines how much noise we want to ignore. Figure 2(b) shows the appearance of an edge map with lower values on the boundary and higher values on the homogeneous regions. As shown in Figure 3, when we scale the pixel values of this edge map to the interval $[0,1]$, it can be used effectively as an image-based term for level set segmentation of masses.

Level set segmentation using a homogeneity edge map

The level set evolution function (3), without the attracting speed term, can be written as:

$$\phi_t + g_I |\nabla \phi| - g_I \varepsilon \kappa |\nabla \phi| = 0, \quad (8)$$

where the first term in (8) provide the expansion speed along the normal direction of the curve, the magnitude is g_I , therefore, in the non-boundary region, we want the edge map g_I to be as large as possible, so that the curve evolution can converge to the real boundary quickly. Ideally, g_I will be zero on the estimated boundary, so that when the curve reaches it, the evolution function becomes $\phi_t = 0$, and reaches equilibrium. But since the gradient $\nabla(G_\sigma * I)$ will never be infinity in the discrete implementation, g_I will never become exactly zero. Thus care and reliable methods are needed to stop the level set evolution when it reaches the estimated boundary.

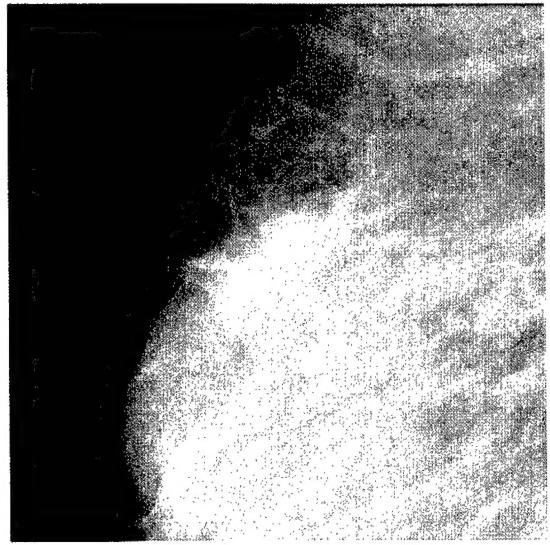
In this research, we used the edge map defined by a homogeneity metric, where the value of the edge map can be zero on a sharp boundary. This provided a more reliable stopping criterion than traditional gradient operators alone. However, because of noise tolerance, a weak boundary could be non-zero. Also, missing boundaries may occur due to the angle of an X-ray projection, therefore, boundary leakage may happen if the boundary definition in the original image is not perfectly clear. This in part motivated our development of the MSAHE algorithm described in the beginning sections in this report. The attraction term (5) introduced in [11] is used to pull back the curve when it passes a boundary. We

added an adaptive indication term that shuts down the expansion speed when the curve became close to the estimated boundary. The evolution function we used was:

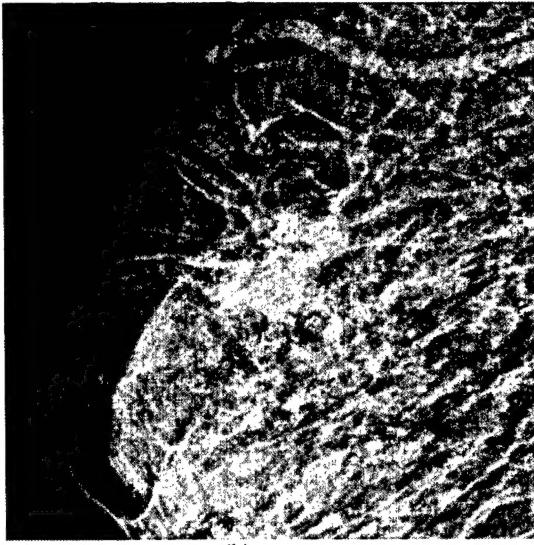
$$\phi_t + g_I \cdot \delta \cdot |\nabla \phi| - \varepsilon \kappa |\nabla \phi| - \beta \nabla P \cdot \nabla \phi = 0. \quad (9)$$

The reason why we took out g_I from the second term is that when the curve approaches the boundary, the value of g_I became very small, and therefore the smoothing effect was eliminated. When using the attracting term, the curve always appeared noisy due to imperfect boundaries in the edge map. This strategy is analogous to those used in parametric deformable models that always keep a constant weighted elastic internal force.

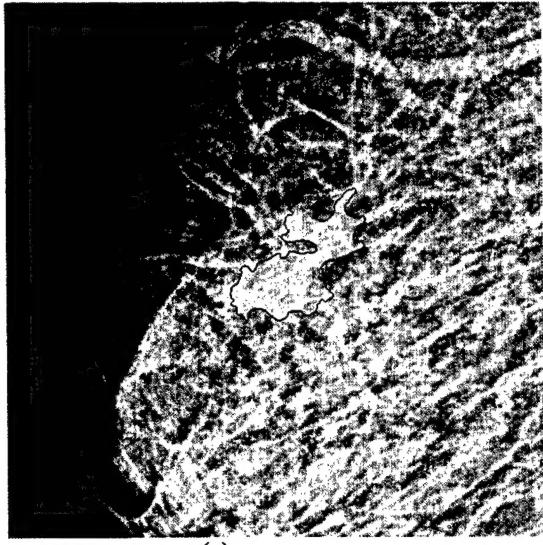
δ is a global indication function such that when the zero level set of ϕ is close enough to the estimated boundary, $\delta = 0$ and otherwise $\delta = 1$. Looking at the computation of the homogeneity map, intuitively the pixel value in the edge map is decreasing from the homogeneous region to the boundary, and reaches the minimum at the boundary. Therefore the pixel value of the edge map actually gives the information of how close a certain pixel is to the boundary. By averaging the values in the edge map over the pixels that represent the zero level set, or finding the maximum value, we can ascertain how close the whole curve is to the target boundary. A threshold is chosen so that when this value is smaller than some threshold $\delta = 0$, and otherwise $\delta = 1$.



(a)



(b)



(c)

Figure 3: Sample mass detection by enhancement within a multi-scale expansion:
(a) original image (b) enhanced image (c) segmentation (shown in red boundaries) based on (b).

3. KEY RESEARCH ACCOMPLISHMENTS

During the past six months of the project we incorporated an expansion based method of detecting masses in digital radiographs.

- We implemented a level-set method of segmentation that made use of a local homogeneity operator for the detection of subtle masses in digital mammography.
- This expansion based method is integrated into our previously described multi-scale expansion framework and will be tested using a local database of digital mammograms as part of a planned validation study.

In our proposal we made the conjecture that existing methods of analysis computed (efficiently) on dyadic scales were not sufficient for the detection of masses: a lesion may be too blurry at one scale, and too precise at the next finer (dyadic) scale. In the first part of this study, we answer the question “Is it sufficient to work with dyadic scales, or is there an absolute need to compute coefficients between the scales?” Indeed, during the course of this investigation we have clearly shown that there is a significant advantage in the capturing the morphology of arbitrary sized masses when using finer “grained” expansions.

Throughout this study we have avoided development of not “reinventing the wheel” whenever possible. We choose to use or modify existing libraries and programs readily available within the mathematical community. For example, we modified under *Matlab* several existing *LastWave* algorithms, including the Discrete Wavelet Transform in two dimensions without downsampling, using the Algorithme à Trous algorithm. An ancillary benefit to this approach is that when this code is made available to the research community it will be easy to use having been built upon “freeware” and commercially available programming environments. All of the programs for computing the expansion and detection algorithms (written in “C” and/or MatLab code) during this the course of the study are available upon request through our web site: bil.bme.columbia.edu”.

We compared in one dimension the CWT and the DWT in order to show a proof of concept concerning any advantage of pursuing refinement of scale. We processed phantom masses, and 1D intensity profiles of real masses mammograms to evaluate feasibility. In order to identify the best scale, we evaluated the use of maxima of the coefficients and a correlated model using three masses of different size. We then evaluated the shape of the “Mexican hat” for suitability in a matched filtering detection paradigm. During the third year we applied enhancement algorithms as a preprocessing step and introduced the notion of “voices” which allowed us to compute representations of masses in between octaves of the expansion.

4. REPORTABLE OUTCOMES

The following publications are presented as reportable outcomes of the investigation. For convenience, we have included selected copies of these manuscripts along with this report.

Publications:

- (0) Y. Jin, A. Laine and C. Imielinska, "An adaptive speed term based on homogeneity for level-set segmentation," *Medical Imaging, Proceedings of SPIE.*, Vol. 4684 (1), pp. 383-390, Feb. 2002, San Diego CA.
- (1) M. A. Birgen, S. Smith, A. F. Laine, "Detection of Masses in Mammography Through Redundant Expansions of Scale," *Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Istanbul, Turkey, October; 2001.
- (2) R. Mekle, A. Laine, S. Smith " Evaluation of a Multi-Scale Enhancement Protocol for Digital Mammography," *Image-Processing Techniques For Tumor Detection*, R. N. Strickland, Ed., Marcel Dekker, New York, NY, (2001) pp. 155-186.
- (3) W. Huda, Y. Jin and A. Laine, "Evaluation of Contrast Enhancement by Digital Equalization in Digital Mammography," *World Congress on Medical Physics and Biomedical Engineering*, Chicago, 2000.
- (4) R. Mekle, A. Laine, S. Smith, C. Singer, T. Koenigsberg and M. Brown, " Evaluation of a Multiscale Enhancement Protocol for Digital Mammography," in *Wavelet Applications in Signal and Image Processing VIII*, A. Aldroubi, A. F. Laine, M. A. Unser, Eds., *Proceedings of the SPIE* Vol. 4119, pp. 1038-1049, San Diego, 2000.
- (5) Koren, A. Laine, S. Smith, E. Nickoloff and F. Taylor, "Visualization of Mammography via Fusion of Enhanced Features," in M. Doi (Editor), *First International Workshop on Computer-Aided Diagnosis*, Elsevier Science, Amsterdam, 1999, pp. 287-303.

List of Personnel:

Andrew Laine, D.Sc. (Principal Investigator)
Suzanne Smith, M.D. (Director of Breast Imaging)
C. Singer, M.D. (Radiologist, Mammography Reader)
T. Koenigsberg, M.D. (Radiologist, Mammography Reader)
M. Brown, M.D. (Radiologist, Mammography Reader)
Edward Nickoloff, Sc.D. (Chief Medical Physicist)
Ralf Mekle, Ph.D. (Graduate Research Assistant)
Yinpeng Jin, M.S. (Graduate Research Assistant)
Mehmet. A. Birgen. (Graduate Research Assistant)

5. CONCLUSIONS

The major problem of level set segmentation is boundary leakage when weak boundaries or occluded parts of a mass boundary are evident. In this research, we kept the expansion term initially for “pushing” the model towards the boundary, and therefore had no need for a restricted initialization. When the expansion term was shut down, we observed that it prevented boundary leakage.

The definition of our homogeneity measurement did not include any prior knowledge of the dimensions, and therefore can be easily extended to higher dimensions and it is computationally efficient because no filtering/convolution is involved. In addition, to collect all the pixels around the zero level set needed to compute the indication function δ was a byproduct of the procedure for constructing the extension speed when iterating the level set function. Therefore this test can be done whenever the construction of the extension speed is performed without extra computation.

Since this research has presented a new edge map, other methods based on speed terms driven by edge maps [16, 20] can be designed to achieve more reliable level set evolution and segmentation. In the future, extensions to higher dimensional datasets and building other desirable speed terms based on “hyper edge maps” and other information of the underlining image are possible.

Overall Summary

The idea of this “Idea Award” was to detect subtle masses in mammograms by tuning the central frequency and width of a basis function that generates overcomplete expansions. By modeling the shape of a mass through this flexibility we hoped to detect small and subtle masses in dense breasts and improve the chances of early detection in screening mammography. In the first part of our investigation, we evaluated existing tools to compute overcomplete expansions of multiscale signals. We compared in one dimension the CWT and the DWT for a proof of concept concerning any advantage of pursuing refinement of scale. We processed phantom masses, and 1D intensity profiles of real masses mammograms to evaluate feasibility. In order to identify the best scale, we evaluated the use of maxima singularities and a correlated model using three masses of different size. Our study answered the question of whether or not dyadic scales were sufficient to detect masses in a dense mammograms. We clearly showed that reasonable approximations of mass shapes could be obtained through overcomplete expansions that computed voices between the traditional dyadic scales.

Our study of one dimension cases answered the question of whether or not dyadic scales were sufficient to detect masses in a dense mammograms. We showed that reasonable approximations of mass shapes could be obtained through overcomplete expansions of a continuous wavelet transform that computed voices between the traditional dyadic scales.

We observed in mathematical phantoms and real masses that a correlation method (between a model of a mass and the values of the computed coefficients) gave approximately the same results when compared to the maxima method (maximum of the coefficients at each scale). We developed a simple scheme to detect masses using these representations. This method based on geometric properties of segmented masses within each expansion was shown to be remarkably stable.

6. REFERENCES

- [1] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: A survey," *Medical Image Analysis*, vol. 1, pp. 91-108, 1996.
- [2] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, pp. 321-331, 1987.
- [3] D. N. Metaxas, *Physics-based deformable models. Applications to computer vision, graphics and medical imaging*, 1997.
- [4] C. Xu, D. Pham, and J. Prince, "Image segmentation using deformable models.," in *Handbook of Medical Imaging*, vol. 2: SPIE, 2000, pp. 129-174.
- [5] T. McInerney and D. Terzopoulos, "Topologically adaptable snakes," presented at 5th International Conference on Computer Vision, 1995.
- [6] R. Durikovic, K. Kaneda, and H. Yamashita, "Dynamic contour: a texture approach and contour operations," *The Visual Computer*, vol. 11, pp. 277-289, 1995.
- [7] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations," *Journal of Computational Physics*, vol. 79, pp. 12-49, 1988.
- [8] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, pp. 158-175, 1995.
- [9] J. Sethian, *Level set methods and fast marching methods*, 1999.
- [10] L. D. Cohen, "On active contour models and balloons," presented at Computer Vision, Graphics, and Image Processing: Image Understanding, 1991.
- [11] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," presented at ICCV'95, 1995.
- [12] M. Tabb and N. Ahuja, "Multiscal Image segmentation by integrated edge and region detection," *IEEE Transactions on Image Processing*, vol. 6, pp. 642-655, 1997.
- [13] P. Saha and J. Udupa, "Scale-based fuzzy connectivity: a novel image segmentation methodology and its validation," presented at SPIE conference on Medical Imaging, 1999.
- [14] P. Saha, J. Udupa, and D. Odhner, "Scale-based fuzzy connected image segmenation: theory, algorithms, and validation," *Computer Vision and Image Understanding*, vol. 77, pp. 145-174, 2000.
- [15] K. R. Castleman, *Digital Image Processing*. Upper Saddle River, NJ: Prentice-Hall Inc., 1996.
- [16] K. Siddiqi, Y. B. Lauziere, A. Tannenbaum, and S. W. Zucker, "Area and Length Minimizing Flows for Shape Segmentation," *IEEE Transactions on Image Processing*, vol. 7, pp. 433-443, 1998.
- [17] A. Yezzi, A. Tsai, and A. Willsky, "A statistical approach to image segmentation for biomodal and trimodal imagery.," presented at ICCV, 1999.
- [18] N. Paragios and R. Deriche, "Geodesic active contours and level sets for the detection and tracking of moving objects," *IEEE Transactions on Pattern Recognition and Machine Intelligence.*, vol. 22, pp. 1-15, 2000.
- [19] A. Sarti and R. Malladi, "A Geometric Level Set Model for Ultrasound Analysis," University of California, Berkeley LBNL-44442, 1999.
- [20] C. Xu and J. L. Prince, "Snakes, Shapes and Gradient Vector Flow," *IEEE Transaction on Image Processing*, pp. 359-369, 1998.

DETECTION OF MASSES IN MAMMOGRAPHY THROUGH REDUNDANT EXPANSIONS OF SCALE

M. A. Birgen¹, S. Smith², A. F. Laine¹

¹Department of Biomedical Engineering, Columbia University, New York, NY, USA

²Breast Imaging Center, New York Presbyterian Hospital, Columbia Presbyterian Medical Center, New York, NY, USA

ABSTRACT

We show that dyadic scales may not be sufficient for the detection of masses in mammograms: a lesion may be too blurred on one scale, and then too fragmented at the next. In this paper, we report on the preliminary evidence of our study using a continuous wavelet transform in two dimensions with arbitrary positioning of a wavelet's center frequency channel tuned to the mass detection problem. Our goal is to detect masses in dense mammograms whose diameter is smaller than 1 cm. The aim is to be able to find the scale where the mass is best represented in terms of analysis.

1. INTRODUCTION

An initial study in one dimension helped us observe that dyadic scales are often not sufficient to detect a mass in a dense mammogram [3]. Below we show this by a continuous wavelet transform, which computes the decomposition on voices between traditional dyadic scales.

2. METHODOLOGY

2.1. Voices and octave

It is possible to expand a signal more finely and compute scales between octaves of traditional dyadic expansions by voices [1]. A voice constitutes a subdivision of an octave. If we consider a wavelet mother ψ , a family of wavelets $\psi_{m,n}(x) = a_0^{-\frac{m}{2}} \psi(a_0^{-m}x - nb_0)$ where a_0 is the dilatation parameter, b_0 is the translation parameter, $(m, n) \in \mathbb{Z}^2$ are possible. In the dyadic case, $a_0=2$ and $b_0=1$, $\psi_{m,n}(x) = 2^{-\frac{m}{2}} \psi(2^{-m}x - n)$. Decomposing N voices per octave means creating N functions $\psi_{n,m}^\nu$ and computing the frame $\{\psi_{m,n}^\nu : (m, n) \in \mathbb{Z}^2, \nu = 1, \dots, N\}$. Analyzing with N

voices means finding N different frequency channels, which correspond to the N frequency localizations of ψ^1, \dots, ψ^N [2], all translated by the same step (Fig. 1b). Such a lattice can be viewed as the superposition of N different lattices of the type shown in Fig. 1a, stretched by fixed amounts in frequency. For example a possible choice for ψ^ν is $\psi^\nu(x) = 2^{-\frac{\nu-1}{N}} \psi(2^{-\frac{\nu-1}{N}} x)$.

If $|\hat{\psi}(\xi)|$, which we assume to be even, peaks around $\pm\omega_0$, then $|\hat{\psi}^j|$ will be concentrated around $\pm 2^{-\frac{j-1}{N}} \omega_0$ in the same way as in the dyadic case. If $|\hat{\psi}|$ has two peaks in frequency at $\pm\xi_0$, $|\hat{\psi}_{m,n}(\xi)|$ then peaks at $\pm 2^{-m} \xi_0$ which are two localization centers of $\psi_{m,n}$.

The equation computing the scale for source given "octave", "current voice" and "number of voices" is $scale = 2^{\frac{octave + current_voice}{number_voices}}$ [3].

Moreover, we adopt the following convention: the first octave (octave number zero) corresponds to the

width between scales $1 + 2^{\frac{1}{number_voices}}$ and 2. The dyadic scale of an octave is the last voice of the octave (scale = $2^{\frac{octave+1}{2}}$). In Fig. 2, we consider a signal of 512 points (2^9). This means 9 octaves (octave 0 to octave 8). The coarsest scale is 512, the finest is $1 + 2^{\frac{1}{number_voices}}$. For example, when we display a second voice of the fourth octave (four voices per octave computed), we obtain the scale $2^{\frac{4+2}{4}}$, that is to say scale 23.

2.2. One-dimensional experiment

We applied programs from libraries in LastWave and Matlab, using a continuous wavelet transform and a discrete wavelet transform. LastWave is a wavelet

signal and image-processing environment, written in C [4]. Wavelab is an extension of Matlab. For the CWT, we concentrated on the first and the second derivative of a gaussian function (Mexican Hat wavelet). We processed phantom signals with three masses of distinct sizes using gaussian additive noise.

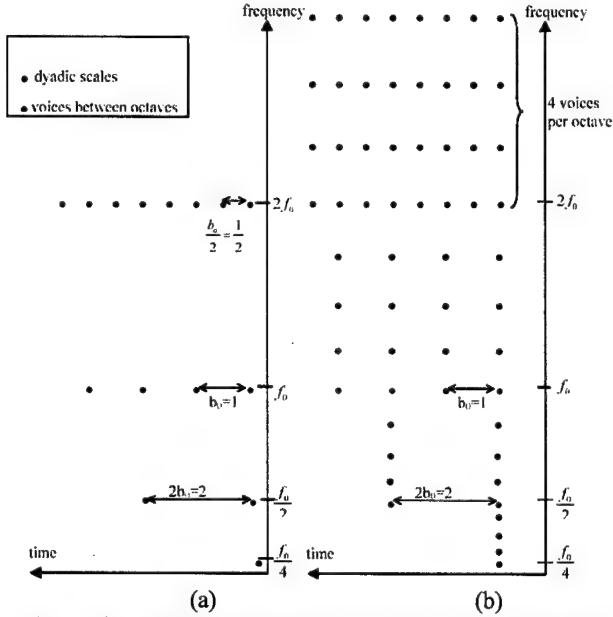


Fig. 1. The time-frequency lattice: (a) for the dyadic wavelet transform, $\psi_{m,n}$ is localized around $2^m b_0$; $a_0=2$ and we assume $b_0=1$; (b) for a scheme with four voices, the different voice wavelets ψ^1, \dots, ψ^4 are assumed to be dilatations of a single

$$\text{function } \psi, \psi^j(x) = 2^{-\frac{j-1}{4}} \psi(2^{-\frac{j-1}{4}} x).$$

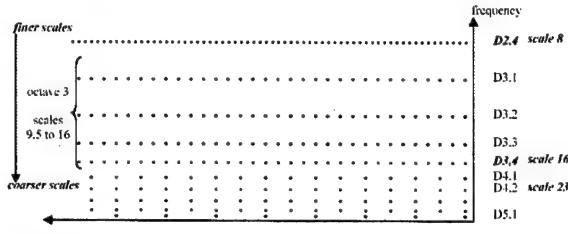


Fig. 2. The time-frequency lattice for a scheme with four voices per octave, including the scale axis.

We plotted two scan line profiles (Fig. 4) of a real mammogram (Fig. 3).

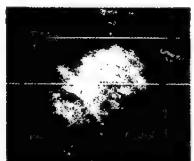


Fig. 3. Real mass from a mammogram. The white lines show the locations of the extracted profiles corresponding to Fig. 4.

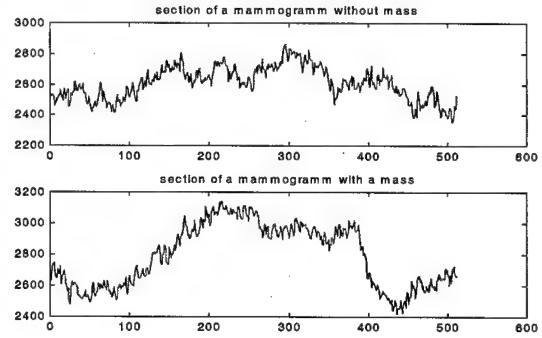


Fig. 4. 1D sections of a real mammogram.

As shown in Fig. 4, we added gaussian noise on the phantom mass so that the 1D signal had approximately the same shape as a real mass. Fig. 5 shows this representation.

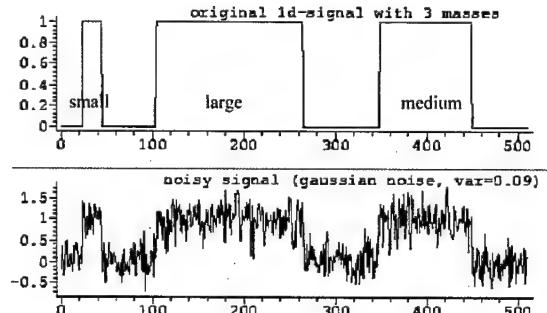


Fig. 5. Phantom signal with an added white gaussian noise of variance 0.1.

Fig. 6 depicts the results obtained without downsampling. The signal was composed of masses with a white gaussian noise of variance 0.1. The wavelet was a Mexican Hat.

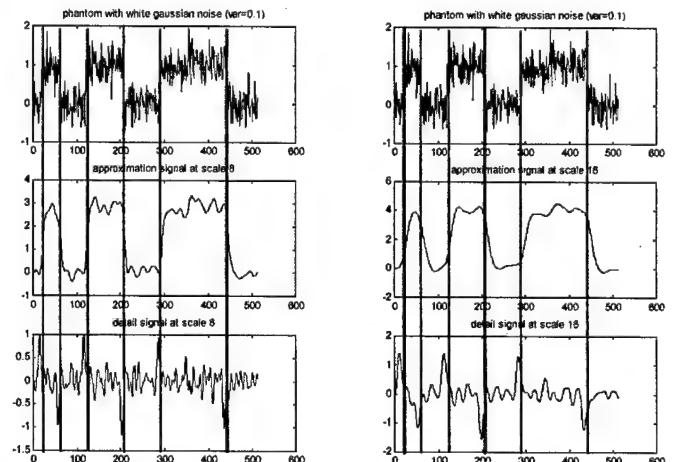


Fig. 6. Analysis by a DWT on a phantom with gaussian noise (var = 0.1). We show approximation and detail signals at scales 8 (left) and 16 (right).

2.3. The 2D CWT

We began the 2D study with phantom masses of white objects on a black background with the addition of white gaussian noise of variance 4. We applied a bias to the magnitude values to preserve the waveform shape and make the signal purely positive [5]. Next, we performed the analysis on a cancerous mass from a mammogram (Fig. 3). We show the biased unthresholded results in Fig. 7, and the thresholded values in Fig. 8.

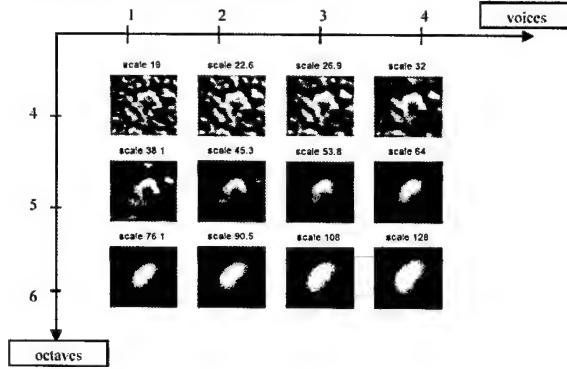


Fig. 7. CWT_2D at octaves 4 to 6, four voices per octave. No thresholding on biased coefficients.

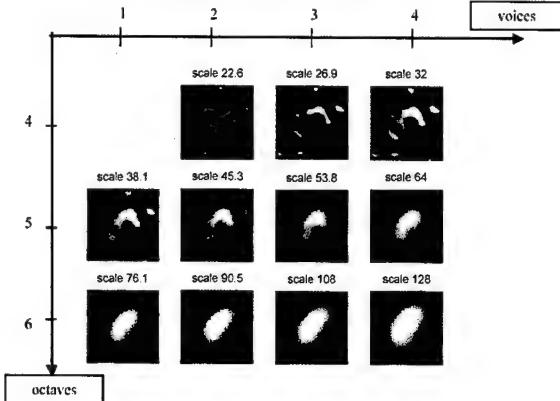


Fig. 8. CWT_2D at octaves 4, 5 and 6, four voices per octave. Coefficients are biased and thresholded among scales (10 for scale 38 to 20 for scale 128).

2.4. Fractional Splines

We have more recently focused on the Fractional Spline Wavelet Transform [6,7]. We have extended the implementation to two dimensions, which was described originally by M. Unser and T. Blu [8]. We used orthonormal filters to compute the details (horizontal, vertical and diagonal) and the approximation coefficients of the image by applying the filters,

$$H_{\perp}^{\alpha}(e^{j\omega}) = \sqrt{2} \left(\frac{1+e^{-j\omega}}{2} \right)^{\alpha+1} \sqrt{\frac{A^{2\alpha+1}(e^{j\omega})}{A^{2\alpha+1}(e^{2j\omega})}}$$

$$G_{\perp}^{\alpha}(e^{j\omega}) = e^{-j\omega} H_{\perp}^{\alpha}(-e^{-j\omega})$$

where $A^{\alpha}(z)$ is the autocorrelation filter of degree α .

The transform is computed for a real mass along the scales for different values of the spline parameter α (Fig. 9).

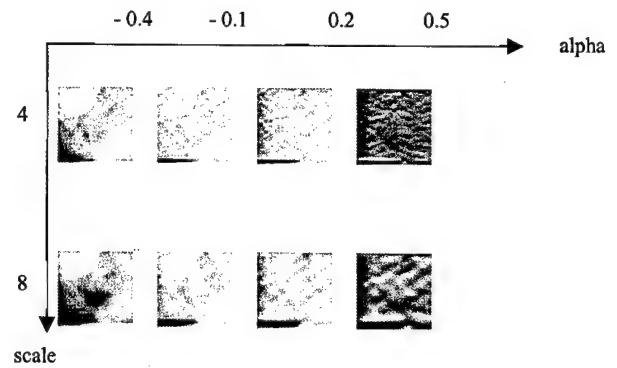


Fig. 9. DWT in 2D at scale 4 and 8, for 4 values of α .

As shown in Fig. 9, we do not always observe a good representation for different values of the parameter α . However, we clearly observe that the detection is better for $\alpha=0.2$. The parameter of the spline is continuous ($\alpha>-0.5$). Therefore, it is interesting to make the parameter vary in order to find the best basis, which suits well a given mass size. However the present transform is only computed at dyadic scales. With a continuous analysis, which would allow decomposition on voices between these scales, we may obtain a richer parameter space so as to identify a best basis for mass detection.

3. Correlation Analysis

Given the results in one dimension, we then implemented a 2D continuous wavelet transform. Our goal was to now find the most suitable scale to detect a mass of arbitrary size. To find the best scale, we displayed the maxima of the coefficients along scales, the third dimension giving the magnitude of the maxima at each scale. In addition, we plotted the correlation between the original mass and the coefficients of the CWT at each scale. We expected to find different optimal scales according to the size of a mass. We tested this by carrying out our algorithm on three different size masses. We first computed for each mass the CWT in 2D on 9 possible octaves (3 voices per octave). Then for each octave and scale we plotted

the maxima of the coefficients of the wavelet decomposition as shown in Fig. 10.

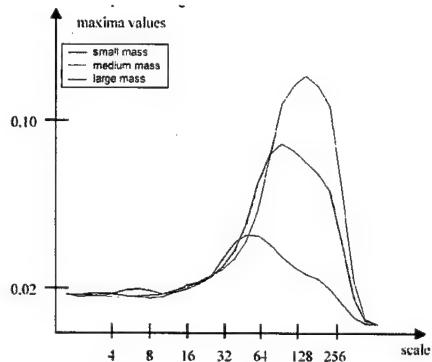


Fig. 10. Evolution of the maxima of the cwt2d across scales.

The positions of the maxima of the decomposition were at scales 40, 81 and 128 for small, medium and large masses respectively.

Next, we performed the CWT in 2D on the same number of octaves and voices. For each scale we calculated the correlation between the original image without noise and the 2D CWT decomposition as shown in Fig. 11.

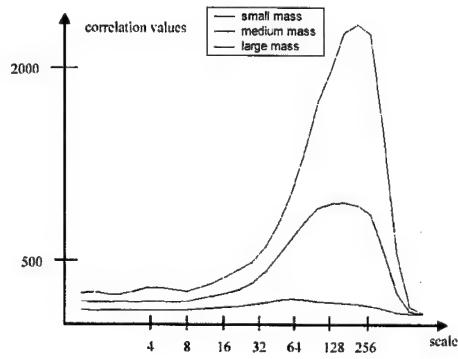


Fig. 11. Correlation between the original image and the biased values of the decomposition.

The positions of the maxima of the correlation were at scales 64, 102 and 161 for small, medium and large masses respectively.

The most suitable scale using the method of the maxima evolution was not the same as the scale identified with correlation. Next, we attempted to find the best scale for a real mass. This time, the best scale to detect the real mass was the same for both methods (maxima evolution and correlation) at scale 161. We also considered a very noisy signal (variance 4), for robustness. We analyzed the maxima of the coefficients and the correlation for different noise settings. From these results we observed that both methods identified same scale values regardless of the amount of added noise.

4. Multi-scale Adaptive Histogram Equalization

Analog and digital mammography often contains 12 bits or more of significant contrast information. Anatomical tissues may occupy significantly different dynamic ranges on display due to difference of X-ray attenuation. By comparison, the human visual system can only perceive less than 100 different gray levels [9]. Thus, contrast enhancement is usually needed for clinical readings. This section discusses one approach for contrast enhancement utilizing multi-scale analysis. Sub-band coefficients were modified by the method of adaptive histogram equalization. To achieve optimal contrast enhancement, the sizes of sub-regions were chosen with consideration to the support of the analysis filters. The enhanced images provided subtle details of tissues that are only visible with tedious contrast/brightness windowing methods currently used in clinical reading.

By properly selecting the decomposition filters, desired features of an object can be separated from noise. Therefore we can selectively enhance features of interest by modifying corresponding components in the transform domain. We used the quadratic spline wavelet function $\psi(x)$, which has compact support and is continuously differentiable. It is the derivative of a cubic spline function $\theta(x)$ as seen in Fig. 12. It can be shown that by using a wavelet that is the derivative of a smoothing function the wavelet transform $W_2^d f$ of the signal f is proportional to the derivative of the signal smoothed at scale 2^j . The wavelet transform can then be considered as an adaptive (scale dependent) detection procedure that finds signal variation points in two orthogonal directions x and y [10].

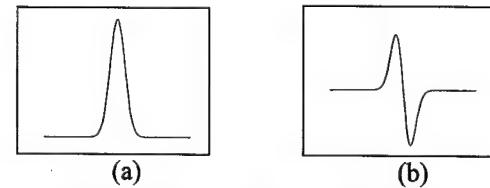


Fig 12. (a) Cubic spline smoothing function $\theta(x)$. (b) Quadratic spline wavelet $\psi(x)$ of compact support defined as the derivative of the smoothing function.

In Fig 13, we present results on mammography data, which shows significant improvement over existing traditional window and leveling techniques used in soft-copy stations. The contrast limited adaptive histogram equalization (CLAHE) clearly enhances monographic features.

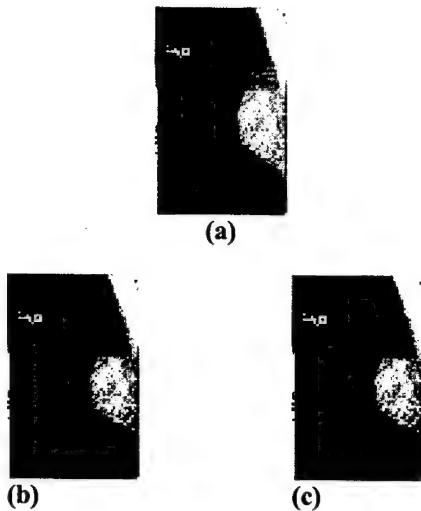


Fig. 13. (a) original film, (b) detailed window & leveling by a radiologist, (c) contrast limited adaptive histogram equalization (CLAHE).

5. ROC Study: Experimental Design

Our study focused on density 3 and 4 mammograms, on BiRads scale. A total of 60 cases, subdivided into 2 groups (15 cancers & 15 normals) were read by three radiologists [11]. The diagnosis for mammograms included BI-RAD (0-5), LOC value (1-5) and localization of detected lesion.

We have divided the radiologists into 2 groups:

Group 1: "Softcopy Display" + Interactive Contrast Enhancement.

Group 2: "Softcopy Display" only.

Our softcopy monitors consist of high-resolution (2048x2560) dual Barco 5MP1H displays. Fig. 15. depicts a radiologist participant in the study along with dual monitor displaying monographic data.

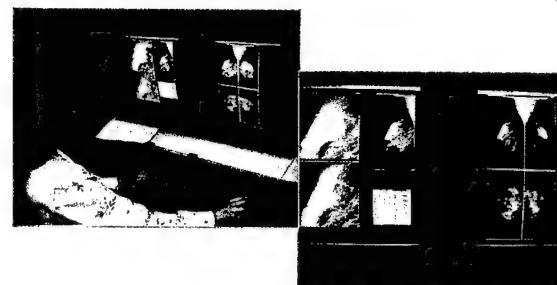


Fig. 15. Actual, experimental study.

The ROC analysis can be seen in Fig. 16. The area under the curve with computer aided enhancement and without enhancement is 0.9136 and 0.8405 respectively. The computer-aided diagnosis brings a noticeable improvement in cancer screening.

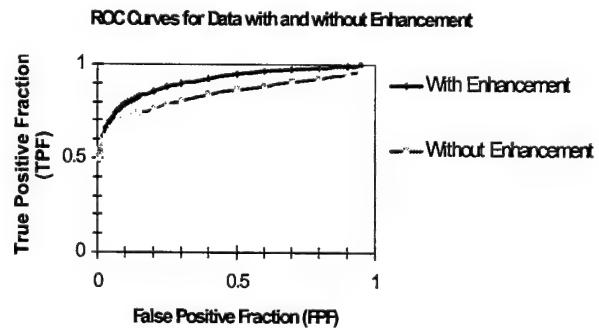


Fig. 16. ROC curves for data from Group 1 ("with Enhancement") and from Group 2 ("without Enhancement") [11].

6. CONCLUSION

Our studies in one and two dimensions suggest that dyadic scales are often not sufficient to detect a mass in a dense mammogram. We showed the advantage of a continuous wavelet transform, which computed an expansion on voices between the common dyadic scales. We saw on real images of masses extracted from digitized mammograms that a correlation method between a known mass and the values of computed coefficients yielded approximately the same results, as a maximum method evolution. Thus, this study suggests that it is possible and of value to tune an analysis between octaves, for the detection of subtle masses in mammograms.

On a second front, a multi-scale adaptive histogram equalization method was reported here, which showed promising results on mammography interpretation. We claimed that the advantage of this method comes from combining the local enhancement ability of AHE, and the selectivity of spatial-frequency components from wavelet analysis. The overall diagnostic sensitivity compares favorable with state-of-art enhancement methods, and also circumvents and reduces some of the artifacts visualized with existing methods. The ability of simultaneously displaying the full dynamic contrast range was shown to be efficient in terms of interpretation time. The diagnostic performance showed the possibility of building a new "Power Windows" scheme for clinical usage.

Following the conventional three-windows settings, we can also tailor the parameters to find the best enhancement for particular abnormalities based on their spatial-frequency properties. This method proves to be of value in isolating cancer masses of diameter 1cm or less. We certainly expect a more reliable diagnosis compared to existing windowing schemes.

5. ACKNOWLEDGMENT

This work was supported in part by U.S. Army Medical Research and Material Command, DAMD17-99-1-9154.

7. REFERENCES

- [1] I. Daubechies, "*Ten Lectures on Wavelets*," Philadelphia, PA: SIAM, 1992.
- [2] A. Laine, W. Huda, D. Chen, and J. Harris, "Segmentation of masses using continuous scale representations", in Digital Mammography '96, Proceedings of the 3rd International Workshop on Digital Mammography, Chicago, U.S.A., 9-12 June 1996, K. Doi, M. L. Giger, R. M. Nishikawa, and R. A. Schmidt, Editors, Amsterdam, The Netherlands: Elsevier, 1996, pp. 447-450.
- [3] D. Chen, J. G. Harris, and A. Laine, "Automatic scale detection", in Visual Communications and Image Processing, Proceedings of SPIE, Orlando, FL, March 17-20, 1996, vol. 2727, pp. 960-972.
- [4] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, CA: Academic Press, 1998.
- [5] A. F. Laine, S. Schuler, J. Fan, and W. Huda, "Mammographic feature enhancement by multiscale analysis", IEEE Transactions on Medical Imaging, vol. 13, no. 4, pp. 725-740, December 1994.
- [6] M. Unser, T. Blu, "Construction of fractional spline wavelet bases", Proc. SPIE "Wavelets applications in Signal and Image Processing VII", Denver, CO, 1999, vol. 3813.
- [7] M. Unser, T. Blu, "Fractional splines and wavelet", SIAM Review, 1999.
- [8] M. Unser, T. Blu, "The fractional spline wavelet transform: Definition and implementation", Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on, Volume: 1, 2000.
- [9] J. B. Zimmerman, S. B. Cousins, K. M. Hartzell, M. E. Frisse, and M. G. Kahn, "A Psychophysical Comparison of Two Methods for Adaptive Histogram Equalization", Journal of Digital Imaging, vol. 2, pp. 82-91, 1989.
- [10] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 14, pp. 710-732, 1992.
- [11] R. Mekle, A. Laine, S. Smith, "Evaluation of a Multi-Scale Enhancement Protocol for Digital Mammography", in Image-Processing Techniques For Tumor Detection, R. N. Strickland, Ed., Marcel Dekker, New York, NY, (2001) pp. 155-186.

Evaluation of a Multi-Scale Enhancement Protocol for Digital Mammography.

Ralf Mekle^a, Andrew Laine^{a, b}, and Suzanne Smith^b

^a Department of Biomedical Engineering, Columbia University

^b Department of Radiology, Columbia-Presbyterian Medical Center

ABSTRACT

We have carried out a receiver operating characteristics (ROC) study for the enhancement of mammographic features in digitized mammograms. The study evaluated the benefits of multi-scale enhancement methods in terms of diagnostic performance of radiologists. The enhancement protocol relied on multi-scale expansions and non-linear enhancement functions. Dyadic spline wavelet functions (first derivative of a cubic spline) were used together with a sigmoidal non-linear enhancement function [1], [2]. We designed a computer interface on a softcopy display and performed an ROC study with three radiologists, who specialized in mammography. Clinical cases were obtained from a national mammography database of digitized radiographs prepared by the University of South Florida (USF) and Harvard Medical School. Our study focused on dense mammograms, i.e. mammograms of density 3 and 4 on the American College of Radiology (ACR) breast density rating, which are the most difficult cases in screening, were selected. To compare the performance of radiologists with and without using multi-scale enhancement, two groups of 30 cases each were diagnosed. Each group contained 15 cases of cancerous and 15 cases of normal mammograms. Conventional ROC analysis was applied, and the resulting ROC curves indicated improved diagnostic performance when radiologists used multi-scale non-linear enhancement.

Keywords: Multi-scale analysis, ROC analysis, contrast enhancement, digital mammography, softcopy display.

1. INTRODUCTION

Recently, research has focused on the development of digital displays and softcopy workstations for digital mammography. Limited spatial resolution, luminance, and dynamic range cannot be solved simply by hardware improvements or computer programming alone. A possible solution of these problems is the application of multi-scale contrast enhancement techniques derived from non-linear models.

Radiologists are mostly familiar with films where the Modulation Transfer Function (MTF) is approximately equal to 2^8 gray levels of contrast resolution. However, images acquired with digital detectors can record at least 2^{12} different gray levels of intensity and are now commercially available. The wealth of dynamic range within these digital acquisition systems provides strong evidence that the signal-to-noise-ratio (SNR) can be increased in digital mammography. For expert radiologists the human visual system can detect at most 2^7 shades of gray. These considerations motivate the need for judicious methods of processing of digital radiographs that can optimize the bandwidth of the human visual system. We have designed enhancement software that is well adapted for this purpose and provides a "data mining" tool to map and make visible selected "quantum levels" of information living within the wide range of contrast resolution provided by digital detectors.

Medical imaging is a field in which quantitative accuracy and qualitative fidelity are paramount. In any image enhancement process distortion of the original image and artifacts are not affordable. Multidimensional feature enhancement via wavelet analysis has been previously demonstrated on mammograms [3], [4], [5], [6], [7], [8] and is a powerful tool for processing digital medical images without artifacts. The enhancement process adjusts multi-scale coefficients at some particular spatial-frequency scale by increasing, decreasing or resetting their values. Each image is then reconstructed with modified coefficients. This simple enhancement technique relies on the idea that features of interest in a given radiograph are detectable at a particular scale and can be amplified, whereas noise and less clinically interesting features may live at other levels of analysis whose visual appearance can be diminished or eliminated in a reconstructed image. Further results and detailed descriptions of these methods can be found in [9], [10], [11], [12], [13], [14], [15].

Surprisingly, there have been very few studies carried out to evaluate the benefits of multi-scale enhancement methods in terms of diagnostic performance. Our study aimed at providing quantitative evidence of these benefits. ROC analysis [16] is most commonly used in medical imaging for such purposes, though alternative statistical approaches can be found as well [17]. ROC curves have been compared to evaluate the visibility of malignancies [18], mass detection techniques [19] or algorithms for computer-aided diagnosis (CAD) that use neural networks [20].

The chapter is organized as follows. In Section 2 we describe a protocol for multi-scale non-linear contrast enhancement. After a short overview of the use of multi-scale expansions for contrast enhancement we discuss the dyadic spline wavelet selected, its implementation, and how a non-linear enhancement function is applied to multi-scale coefficients. Section 3 addresses the design of a graphical user interface (GUI) that was developed to carry out the ROC study including high-performance displays and specialized hardware for softcopy display of digital mammograms. Next, the ROC study itself together with its results and subsequent data analysis is presented in Section 4. After a discussion of the results of the study, conclusions and possible directions of future research are presented in Section 5.

2. ENHANCEMENT PROTOCOL

2.1. Contrast Enhancement via Multi-scale Expansions: A Short Overview

We summarize below, the advantages of the use of overcomplete multi-scale representations for adaptive contrast enhancement of digital mammograms. Critically sampled multi-scale representations are not suitable for detection and enhancement tasks because of aliasing effects introduced during downsampling of the analysis [21], [22]. However, overcomplete representations avoid such aliasing artifacts and have the desirable property of being shift invariant [23], [24]. Indeed, this property ensures that the spatial locations of any mammographic finding within in an image are preserved across all scales. Thus, in our approach the transform coefficient matrix size at each scale remains the same as the original spatial resolution of the digital mammogram, since there is no downsampling across each level of analysis.

Overcomplete multi-scale analysis and reconstruction algorithms using dyadic scales previously developed in [25], [26], and [27] were used as an initial choice of analysis function for our enhancement protocol. The implementation was carried out using several lowpass and highpass filters with localized frequency support. At each level of the multi-scale expansion an input image is decomposed into a coarse approximation and detailed structures. The coarse approximation is the output from applying a lowpass filter, and the detailed structures are obtained from highpass filtering. The approximation image corresponds to scaling coefficients, whereas the details extracted from the approximation are wavelet coefficients at a particular scale. This procedure is successively repeated on the approximation image to obtain multiple levels of analysis. The coarsest approximation is often referred to as “dc-cap”. A gain or enhancement function modifies the matrices of coefficients that have been isolated by the filters at each level and may boost coefficients at some scales and/or attenuate others. If the filters meet a perfect reconstruction condition, the image can be reconstructed from its wavelet representation of scaling and wavelet coefficients [28]. The filter bank implementation of enhancement processing by an expansion-reconstruction algorithm for 2 levels of analysis is schematically illustrated in Figure 1. Image reconstruction that is also accomplished by appropriate filtering operations is presented in a simplified manner in Figure 1.

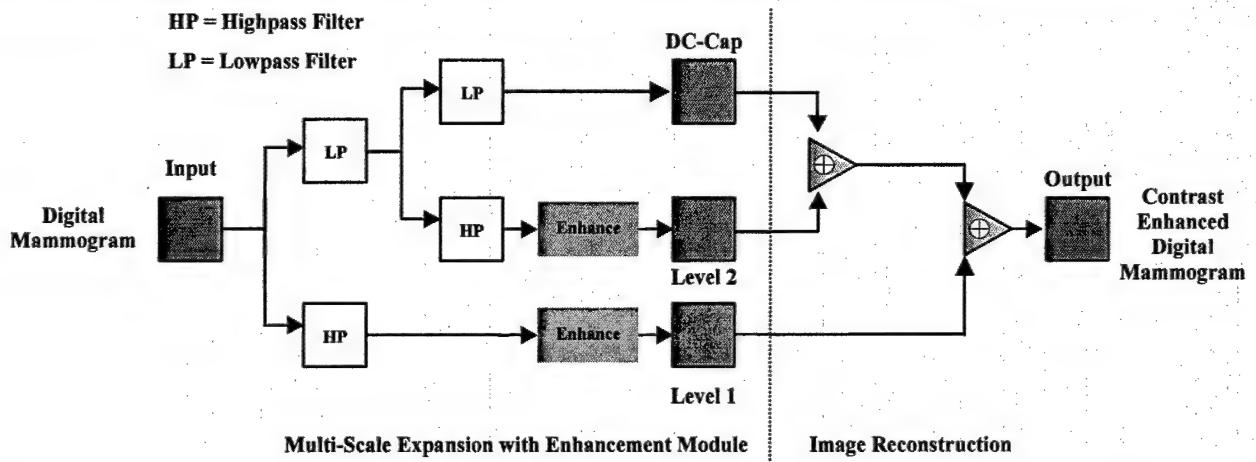


Figure 1: Multi-scale analysis with non-linear contrast enhancement: Schematic of filter bank implementation. In the left part multi-scale expansion with enhancement for 2 levels of analysis is shown, and reconstruction is presented (in a simplified manner) in the right part.

The modified matrices of coefficients are simply “plugged in” during reconstruction producing a “focused” subband enhancement. As shown above, the enhancement function can be implemented independently of a particular set of filters and easily incorporated into a filter bank to provide the benefits of multi-scale enhancement [1], [29].

2.2. High Speed Implementation to Support Interactive Processing

Similar to orthogonal and biorthogonal discrete wavelet transforms [30], the discrete dyadic wavelet transform can be implemented within a hierarchical filtering scheme. Let an input signal $x(n)$ be real, $x(n) \in l^1(Z)$, $n \in [0, N-1]$, i.e., $x(n)$ is supported on the index interval $[0, N-1]$, and let $X(\omega)$ be its Fourier transform. Depending on the length of each filter impulse response, filtering an input signal may be computed either by multiplying $X(\omega)$ by the frequency response of a filter or by circularly convolving $x(n)$ with the impulse response of a filter. Of course, such a periodically extended signal may change abruptly at the boundaries and cause artifacts. A common remedy for such a problem is realized by constructing a mirror extended signal

$$x_{me}(n) = \begin{cases} x(-n-1), & \text{if } n \in [-N, -1] \\ x(n), & \text{if } n \in [0, N-1] \end{cases}$$

where we chose the signal $x_{me}(n)$ to be supported in $[-N, N-1]$. In [1] it is shown how a mirror extension is a particularly elegant solution in conjunction with symmetric/anti-symmetric filters, since a signal is of a particular type of symmetry at each stage of the filter bank. The optimized circular convolution described in [1] was implemented in native “ANSI C” to speed up performance for multi-scale decomposition and image reconstruction. Parameters of this algorithm included number of levels of analysis, gain, and threshold. This algorithm was incorporated into a graphical user interface (GUI) developed during the preparation of the study.

As a further goal, we envision developing feature specific enhancement protocols for each type of lesion. An enhancement protocol would consist of a multi-scale expansion of a mammogram by a specific basis and an associated non-linear enhancement function that is best matched to a specific type of lesion, e.g. microcalcifications. For the study under consideration, a dyadic spline wavelet function was used as the basis, and a non-linear sigmoidal function was applied as the enhancement function. Both are described in greater detail next.

2.3. Dyadic Spline Wavelet Algorithm

The wavelet transform of a signal $f(x)$ at scale s and position x is defined by $Wf(u, s) = f * \psi_{u,s} = \int_{-\infty}^{+\infty} f(x) \frac{1}{\sqrt{s}} \psi^*(\frac{x-u}{s}) dx$,

where the function f is projected on a family of translated and dilated basis functions (wavelets) $\psi_{u,s}(x) = \frac{1}{\sqrt{s}} \psi\left(\frac{x-u}{s}\right)$.

$\psi(x)$ is the mother wavelet of zero average. Both, translation and dilation parameters u and s are continuous for the continuous wavelet transform. To allow fast numerical implementation of discrete wavelet transforms, Mallat and Zhong [31] introduced the dyadic wavelet transform, where the scale parameter varies only along the dyadic sequence $\{2^j\}$, with $j \in Z$. Extending this approach to two dimensions by the use of a tensor product yields the 2-D dyadic wavelet transform that partitions plane orientations into two bands. This means that there are two channels of analysis along orthogonal directions x and y . The wavelet transform of the 2-D signal $f(x,y)$ at the scale 2^j has two components defined by:

$W_{2^j}^1 f(x,y) = f * \psi_{2^j}^1(x,y)$ and $W_{2^j}^2 f(x,y) = f * \psi_{2^j}^2(x,y)$, with $\psi_{2^j}^d(x,y) = \frac{1}{2^{2^j}} \psi^d(\frac{x}{2^j}, \frac{y}{2^j})$, ($d=1,2$). We used the quadratic spline wavelet function $\psi(x)$ defined by Mallat and Zhong in [31] of compact support and continuously differentiable. Its

Fourier transform can be derived as $\hat{\psi}(\omega) = (j\omega) \left(\frac{\sin(\omega/4)}{\omega/4} \right)^4$. $\psi(x)$ is the first derivative of a cubic spline smoothing

function $\theta(x)$, whose Fourier transform is $\hat{\theta}(\omega) = \left(\frac{\sin(\omega/4)}{\omega/4} \right)^4$ [1]. These functions are displayed for the one-dimensional case in Figure 2 below.

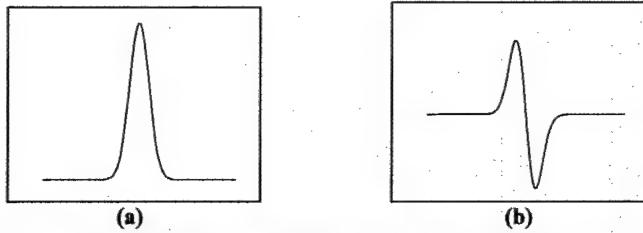


Figure 2: (a) Cubic spline smoothing function $\theta(x)$. (b) Quadratic spline wavelet $y(x)$ of compact support defined as the first derivative of the smoothing function.

Using a wavelet that is the derivative of a smoothing function it can be shown that the wavelet transform $W_{2^j}^d f$ of the signal f is proportional to the derivative of the signal smoothed at the scale 2^j [32]. The coefficients of modulus maxima detection are then equivalent to an adaptive sampling that finds signal variation points in the two orthogonal directions x and y . As images represent finite energy signals measured at some finite resolution, we cannot compute the wavelet transform at scales below the limit set by this resolution. We applied this analysis at dyadic scales varying from 1 (original signal) to the limit imposed by acquisition (digitizer sampling rate). Figure 3 shows an example for one level of an overcomplete wavelet expansion of a region of interest (ROI) with a spiculated mass at a dyadic scale, and in Figure 4 wavelet coefficients of microcalcifications at the finest dyadic scale are presented.

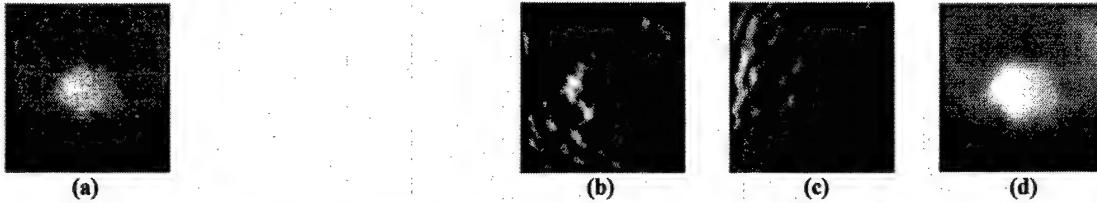


Figure 3: Level 5 of an overcomplete dyadic wavelet expansion of a spiculated mass. (a) Original image. (b) Horizontal details. (c) Vertical details. (d) Approximation image.

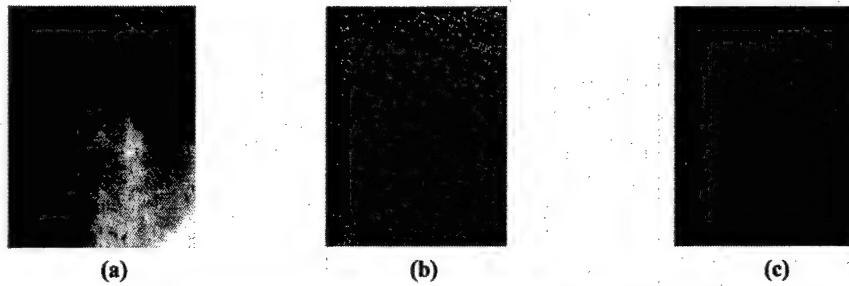


Figure 4: (a) Original ROI with microcalcifications. (b) Horizontal and (c) Vertical dyadic wavelet coefficients.

2.4. Non-Linear Enhancement Function

Modification of selected analysis coefficients within a certain scale can make more obvious indiscernible or barely seen mammographic features [14]. Contrast enhancement was achieved by applying an enhancement function to transform coefficients at selected scales. This operation results in local attenuation or amplification of coefficients. Enhancement or gain functions must be cumulative and monotonically increasing, in order to preserve the order of intensity information in the original image and to avoid artifacts [26]. Figure 5(a) provides a very simple example of a piecewise linear enhancement function. Multi-scale coefficients are denoted w_{ij} , which are modified by applying an enhancement function $f(w_{ij})$. T is the threshold of the function, and α the gain. The effect of the enhancement function depends on the value of the angle θ . For $\theta < 45^\circ$ there is an attenuation of the coefficients ($\alpha < 1$), at $\theta = 45^\circ$ we have the identity function ($\alpha = 1$), and for $\theta > 45^\circ$ there is a smooth amplification of the coefficients ($\alpha > 1$). The values of the two parameters, T and θ (or α), determine the final shape of the enhancement function. Figure 5(b) displays a hard-thresholding function for denoising, where coefficients with modulus $|w_{ij}| \leq T$ are set to zero. Unfortunately, these two particular functions have the disadvantage of being discontinuous at the threshold value $\pm T$. This could result in an abnormal distribution of coefficient values in the output and may create sharp

peaks on both ends of the histogram of a particular output mapping. For this reason, smoother functions, like sigmoids, are preferable and were used in this study. Figure 6 shows an example of such a function as described in [2].

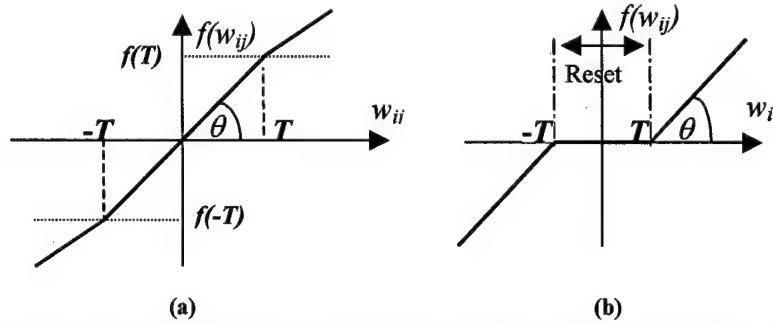


Figure 5: (a) A simple piecewise linear enhancement function, (b) Hard-thresholding function.

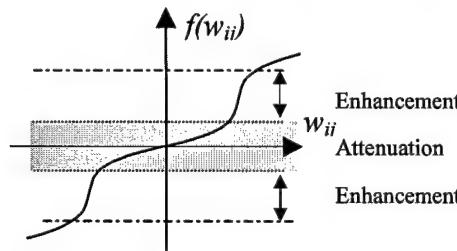


Figure 6: A sigmoidal non-linear enhancement function.

The analytical formulation of the sigmoidal enhancement function as designed in [33], [2] is the following:

$$f(w_{ij}) = a \left[\text{sigm}(c(w_{ij} - b)) - \text{sigm}(-c(w_{ij} + b)) \right]$$

$$a = \frac{1}{\text{sigm}(c(1-b)) - \text{sigm}(-c(1+b))}, \quad 0 < b < 1$$

$$\text{sigm}(y) = \frac{1}{1 + e^{-y}}$$
(1)

Parameters b and c control the threshold and the rate of enhancement (gain) respectively. This enhancement function is continuous, monotonically increasing, and has a continuous first derivative. This ensures that the application of the function will not introduce any new discontinuities of coefficients in the transform domain.

From Figure 6 we see that this enhancement function decreases the value of the coefficients around zero, which is equivalent to a denoising action, while it may increase values of the coefficients outside this range, equivalent to enhancement or amplification. This type of enhancement function, in 'steps', offers a very rich and flexible paradigm to carry out non-linear dynamic analysis of coefficients within a specific scale [34].

There are many criteria for the selection of the enhancement function applied to the coefficients of a particular level of analysis for contrast enhancement. One goal of the study described here was to develop a research tool for testing enhancement functions targeted for specific mammographic features. As this process requires specialized expertise and a substantial time investment, no systematic study of the problem of associating enhancement functions with target features in mammograms has been reported in the literature.

In general, non-linear estimators are signal dependent and behave differently for different realizations of each signal. In this context, Johnstone and Donoho have shown that by considering the signal as deterministic, thresholding of wavelet coefficients gives a nearly optimal estimation of piecewise smooth functions [35], [36]. More specifically, for a noisy signal of size N , thresholding of the wavelet coefficients with $T = \sigma \sqrt{2 \ln(N)}$, where σ is the standard deviation of the coefficients, provides an asymptotically optimal estimator of the original signal in the mini-max sense [36]. Thresholding of wavelet coefficients performs an adaptive smoothing of the image by averaging noisy areas and preserving or enhancing coefficients in areas of sharp transitions. Noise standard deviations can be estimated by determining the median wavelet coefficient value at the finest scale or with local discrete statistical estimation in the transform domain. Using extremely local variances for the

estimation of a threshold leads to a very aggressive posturing of the enhancement function, and represents a high amount of intervention in adjusting the output, while global variance measurements are less noticeable. Superiority of either method depends on the screening protocol used by the radiologist and the kind of analysis to be performed. For example, fine microcalcifications represent high frequency information of the image. We would expect the local variance for such a feature to be high within a selected ROI. Consequently, smooth amplification of coefficients within this particular spatial frequency range (in combination with possibly decreasing the information of other spatial frequencies) will enhance these features of interest. Similar analysis can be done to enhance low spatial frequency features such as masses. A block diagram of the enhancement process for coefficients at selected scales, which are chosen with respect to the particular mammographic feature to be enhanced, is shown in Figure 7 below.

Since the computation of the enhancement parameters uses data dependent information such as local or global coefficient variance, digital and digitized radiographs acquired under different imaging conditions are best processed independently to achieve optimal enhancement. Intrinsic properties of the radiograph are therefore incorporated in the setting of the parameters. In our work we used both coefficient variance computed with respect to a selected ROI and user input (see Section 3.2) to adapt the threshold and gain parameters.

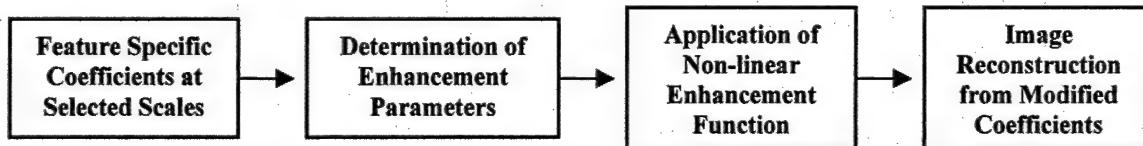


Figure 7: Block diagram of modifying feature specific coefficients at selected scales by applying a non-linear enhancement function.

3. DEVELOPMENT OF A GRAPHICAL USER INTERFACE (GUI)

3.1. Motivation

Running such an enhancement algorithm in a batch mode might be sufficient for single experiments. However, adjustment of parameters tied to a data dependent enhancement function is slow, because of the repeated need to decompose and reconstruct from modified coefficients. A more desirable situation would be to observe the results of modified multi-scale coefficients interactively and to continue the enhancement procedure, until results are visually satisfactory or the decision is made that no further improvement can be achieved. In addition, with introducing fixed enhancement protocols into a clinical screening paradigm, the algorithm must be simple, fast, and user-friendly, i.e. usage of the algorithm should be familiar to the radiologist and intuitive. Since each radiologist may have preferences with respect to contrast in mammograms, it must be possible to adjust parameter settings to individual preferences. Thus, we designed a graphical user interface (GUI) to facilitate carrying out such a study and to create a softcopy display prototype, whose successors might find entrance into clinical screening. We call this application a "test bed" softcopy display tool. Its first version was employed for the ROC study described in the next section.

3.2. Design and Implementation

The graphical user interface (GUI) developed for this study was written in Visual C++ 6.0. The code for the wavelet expansion and image reconstruction that was written in native "ANSI C" to speed up performance could be incorporated and executed in this environment without major modifications, thus shortening development time. Some of the guidelines and considerations for the design and implementation of the GUI are described next.

The prototype interface was primarily designed to process raw 16-bit data. Data was obtained from a national mammography database of digitized radiographs provided by the University of South Florida (USF, "Digital Database for Screening Mammography" (DDSM)). Our database of digitized mammograms (stored on twenty-two 8mm tapes) at the time of the study contained 586 selected cases of malignant lesions, biopsy proven, and 437 cases of normal breasts. More specifically, different types of lesions are represented in the following proportions: 100 round and oval malignant masses, 216 spicular lesions and 248 microcalcifications. 559 cases of dense breasts (density of 3 and 4) with 266 normals and 293 cancerous, referred by radiologists as the most challenging cases, were included in the database.

Images from the mammography database were digitized from film at resolutions of 40 to 50 μm . Image line lengths (# of columns) varied between 2000 and 3000 pixels, and number of rows from 4000 to 5900 pixels. Depending on the scanner utilized for digitization the contrast resolution was either 12 bits or 16 bits per pixel resulting in 15-50 megabytes per view.

To handle this large amount of data and to provide the diagnosing radiologist as much information as possible, all four views (right and left medial-lateral (RMLO, LMLO) and right and left cranial-caudal (RCC, LCC) of a case were loaded into memory and displayed as downsampled images on display screen, which consisted of two high-resolution MegaScan

monitors each with a screen size of 2048 by 2560 pixels. Specialized framebuffers allowed a display of 2^{10} gray levels (see Section 3.3). The four views were aligned to assist the radiologist to look for asymmetries. In addition, one view could be selected, and a viewport could display a selected region of interest (ROI) at full (original) resolution from a selected mammogram. The size of the viewport could be chosen as 512 by 512, 1024 by 1024 or even 2048 by 2048. The center of the ROI was determined through the mouse pointer in a chosen window. Thus, the original mammogram could also be examined through the viewport, if desired. More importantly, suspicious areas could be captured in the viewport and processed through enhancement via the multi-scale expansion described in Section 2. For the enhancement procedure the user could adjust the number of subbands of the expansion as well. After selecting a ROI the image was decomposed onto dyadic wavelet basis functions yielding wavelet coefficients. Coefficients were modified by a sigmoidal non-linear enhancement function, and the image was reconstructed from these modified coefficients in nearly real-time.

Figure 8(a) shows Dr. Koenigsberg, one of three radiologists who participated in this investigation, during the ROC study. Figure 8(b) depicts a typical screen display of the GUI showing additional viewports described above.

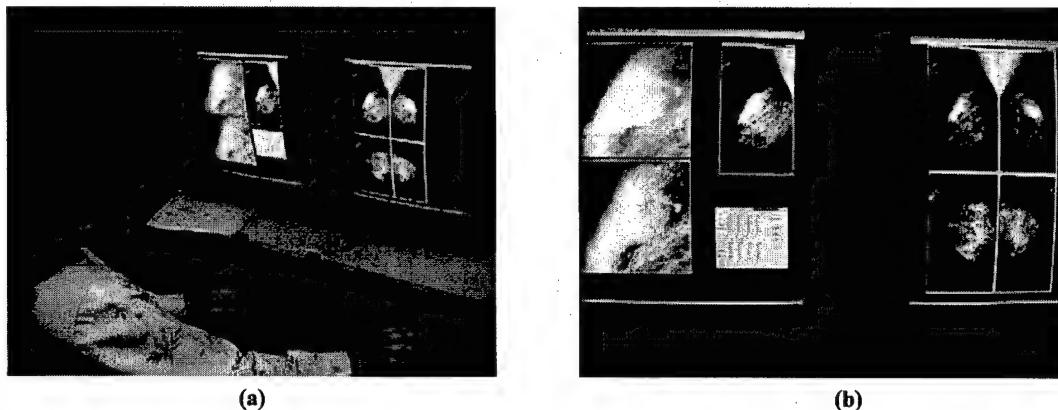


Figure 8: (a) Tova Koenigsberg, M.D., using the GUI during the preliminary ROC study described above. (b) Typical screen display used during the ROC study: four original digitized mammograms of one case on the right monitor, and a selected view, the GUI interface for parameter adjustments, original and enhanced ROI are shown on the left monitor.

As mentioned in Section 2.4 the shape of the enhancement function can be changed through modification of the two parameters gain and threshold. Therefore, each parameter could be adjusted through sliders for each level (subband) of the multi-scale expansion (see Figure 9(b)). On release of the slider button, a reconstruction “event” was “triggered”, and a resulting image presented in an output window. For example, reconstruction of a 512 by 512 matrix for five levels of decomposition (5 subbands) took 5 to 6 seconds. For four subbands reconstruction time shortened to 4 to 5 seconds. Reconstruction times t_{recon} for different sizes of the ROI and different number of levels of analysis are presented in Table 1. However, reconstruction time can certainly be improved to achieve true real-time performance, by employing faster algorithms.

Size of Region of Interest (ROI)	t_{recon} for 4 Levels of Analysis	t_{recon} for 5 Levels of Analysis
512 x 512	4-5 seconds	6-7 seconds
1024 x 1024	19-20 seconds	24-25 seconds

Table 1: Reconstruction times t_{recon} for two different levels of analysis and two sizes of ROI.

After processing, enhanced images could be saved together with information about the location of the ROI (the position of the ROI was marked in its corresponding downsampled view) to facilitate evaluation of a particular diagnosis for each case in comparison with the “ground truth” provided in the USF database. All suspicious areas in a case could be carefully examined by sequentially choosing different views and multiple ROIs.

Figure 9(b) shows the test bed interface as an illustration. Interactive (real-time) enhancement was accomplished via sliders shown in the graphical user interface (GUI). The enhancement operation relied on the optimality of parameters derived from their non-linear models and on the strategy employed for the type of enhancement applied to each subband of coefficients (amplification, preservation or diminution). Selected subband coefficients at a particular level could be strongly suppressed by choosing large thresholds (> 2) and small gains (< 1), which can be desirable for the elimination of (structured and acquisition) noise, or normal benign anatomical (fibroglandular) structures.

Since the size of digital mammograms is quite large, an ROI (fixed at either 512 x 512 or 1024 x 1024) within the original image was chosen to avoid computing over regions that do not contain suspicious areas. This is also shown in Figure 9, where Figure 9(a) exhibits an original digitized mammogram with a 512 x 512 ROI that contains a possible mass.

Figure 9(c) and Figure 9(d) display this ROI before and after enhancement via non-linear modification of multi-scale coefficients, respectively.

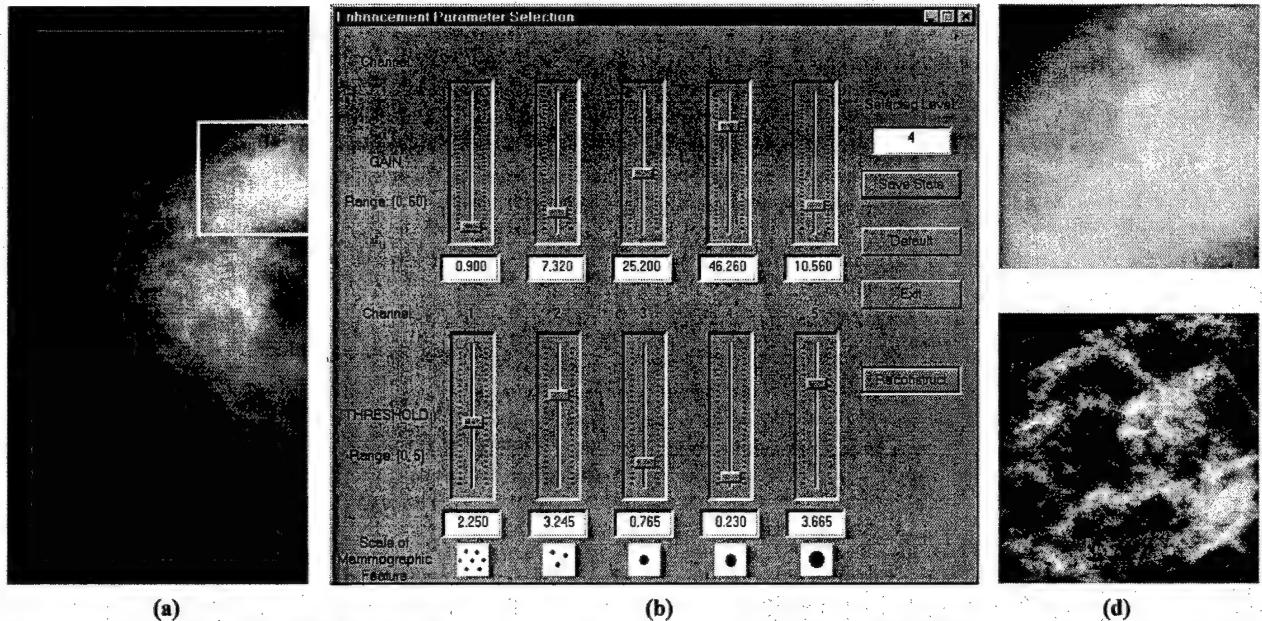


Figure 9: (a) Original mammogram with selected ROI containing a mass, (b) Multi-Scale Contrast Enhancement (MSCE) GUI, (c) Original ROI, and (d) Enhanced ROI.

3.3. Display and Hardware Settings

The enhancement protocol was executed on an IBM IntelliStation Z Pro Professional Workstation Type 6865. This machine had two Intel Pentium II Xeon microprocessors (450 MHz), 512 Mbytes of RAM and was equipped with 36 Gbytes of hard disk space. Windows NT 4.0 with service pack 4 was the operating system.

To explore the richness of information quantized at 16-bit per pixel (bpp) grayscale data (65536 shades of gray), the IBM IntelliStation workstation was equipped with two BARCOMed 5MP1H Graphics controllers. These are high-resolution display subsystems for the PCI bus with a resolution of 2048×2560 pixels each, a digital-to-analog converter (DAC) capable of 1024 shades of gray, and real time window leveling. With the BARCO framebuffers, an extended hardware palette of nearly 16,000 entries could be accessed through specialized "C" function calls that were part of a library provided to us as developers for BARCO/Metheus. Using these library functions, the extended palette was loaded with a ramp of 4096 shades of gray corresponding to 12-bit resolution. Images stored in 16-bit per pixel format, were rescaled to 12 bpp, if necessary (most of the mammograms were digitized at a resolution of 12 bpp), and then displayed at full resolution. Direct access to the video framebuffer also sped up the display process useful for updating and refreshing the different views on the screen.

Two high-resolution MegaScan monitors were attached to this workstation providing dual headed display on a single logical framebuffer or virtual desktop of 4000×2048 pixels, respectively with Windows NT 4.0. To ensure the accurate depiction of the same image quality on both screens, a BARCO P1500 luminance photometer was used. It recognized the 1024 shades of gray displayed by a monitor and had a range of 0-450ft-L. Both monitors were calibrated to correct for non-linearity of display properties through gamma correction.

Lighting conditions were controlled for the ROC study to model reading room conditions. The ambient light intensity was measured with the luminance photometer to be 12.802659 candela/m². It is worthwhile to note that the optimality of enhancement parameters is independent of the CRT display quality and the image acquisition quality. As their computation is data driven, they are adapted to signal content and its characteristics. As our radiologists gave us feedback on the quality of the enhancement, we can adjust these initial default settings in future studies.

4. DESCRIPTION OF THE RECEIVER OPERATING CHARACTERISTICS (ROC) STUDY

The first receiver operating characteristics (ROC) study focused on overcomplete dyadic wavelets for enhancement of mammographic features in digitized mammograms. Specifically, dyadic spline wavelet functions were used together with a sigmoidal non-linear enhancement function explicitly described in Section 2. The ROC study included three radiologists

specialized in mammography. The Director of the Breast Imaging Center at Columbia-Presbyterian Medical Center, Dr. Suzanne Smith, assisted in the selection of cases.

4.1. Selection of Cases

To measure the benefits of diagnosing digitized mammograms with enhancement through multi-scale expansions, we focused on dense mammograms, i.e. mammograms of density 3 and 4 on the American College of Radiology (ACR) breast density rating, which are the most difficult cases in screening. In general, the enhancement protocol aimed at improving the detection and localization of mammographic features, such as microcalcifications, masses, and spicular lesions without introducing "false-positives".

To compare the performance of radiologists with and without using the enhancement tool, two groups of 30 cases each were presented. Each group contained 15 cases of cancerous and 15 cases of normal mammograms. As mentioned above, a national mammography database of the University of South Florida provided "ground truth" (mostly through biopsy) for the selected cases. The selection was carried out very carefully under the guidance of a mammographer (Dr. Smith), in order to find rather challenging cases of similar difficulty for each group. Images showing metal markers ("bibis") to indicate suspicious regions of breast tissue were avoided as well as obvious malignancies. Due to time constraints the number of cases was limited for this initial study.

4.2. Paradigm of Diagnosis of Study

For each case presented to the radiologist, the enhancement procedure followed was the following:

Paradigm A: Without Enhancement:

The radiologist made a diagnosis based only on the four original displays and the viewport. No processing of ROIs was allowed.

Paradigm B: With Enhancement:

The radiologist selected an ROI in one of the views and could apply multi-scale enhancement. Four levels of coefficients were computed. The radiologist then evaluated the quality of an enhanced ROI and adjusted the equalizer sliders of a channel to improve the visual quality of suspicious regions. Once he/she was satisfied with the visual result or if he/she judged that additional benefit could not be achieved, he/she made a diagnostic decision.

A diagnosis included specifying all lesions found and assigning a BI-RAD scale to each breast and the case. In addition, the radiologist was asked to choose a level of confidence (LOC) for each positive diagnosis, i.e. cancer is present, on an integer scale from 1 (definitely negative, i.e. total confidence that there are no malignant lesions) to 5 (definitely positive, i.e. total confidence that there is a malignant lesion). The value for the LOC was used in the analysis of data to decide whether a lesion was classified as malignant or benign (please see discussion of LOC ratings in Section 4.4).

4.3. ROC Data

Table 2 and Table 3 summarize the data acquired during the study. Group 1 comprises the set of cases, where the radiologists were allowed to take advantage of the enhancement protocol, whereas Group 2 contains those cases, where no processing could be applied. Each of the tables shows the case numbers, the case designation and total number (#) of lesions for each case according to the mammography database (DB), and for each of the three mammographers the BI_RAD rating and level of confidence (LOC) values. The BI_RAD rating could be chosen from the standard categories 0-5 with 0 meaning that additional information for a more confident diagnosis was needed. In such cases, the radiologists were asked to also select a BI_RAD rating different from 0, if they were asked to make a diagnosis without any additional information. This number is shown in parentheses for such cases.

In each table both groups are sorted into actually-negative cases (normals with "0" lesions) and actually-positive cases (cancers with, at least "1" malignant lesion), since this is required for subsequent analysis of the data.

Group1 (with Enhancement)			Mammographer 1		Mammographer 2		Mammographer 3	
Case #	Database	DB Total # of Lesions	BI RAD	LOC	BI RAD	LOC	BI RAD	LOC
2	A 0058	0	4	3	1	1	3	2
5	A 0069	0	1	2	1	1	1	1
6	A 0041	0	3	2	1	1	1	1
7	A 0077	0	3	2	2	1	2	1
9	A 0064	0	2	2	2	1	2	2
13	A 0067	0	0(3)	2	1	1	0(3)	3
15	A 0080	0	0(3)	3	2	1	2	1
16	A 0089	0	3	3	1	1	1	2
19	A 0062	0	2	2	1	1	2	1
21	A 0057	0	2	2	1	1	0(3)	3
24	A 0072	0	1	2	1	1	1	1
25	A 0070	0	1	2	0(3)	2	1	2
26	A 0068	0	1	2	1	1	2	1
28	A 0039	0	3	2	1	1	0(4)	3
30	A 0092	0	3	2	1	1	1	1
1	B 3044	1	4	4	4	4	4	3
3	B 3073	1	3	2	3	2	4	3
4	B 3006	1	5	5	5	5	5	5
8	B 3032	1	0(3)	2	5	4	4	4
10	B 3107	1	5	4	4	4	5	4
11	C 0060	1	0(3)	3	0	3	0(4)	3
12	B 3057	1	4	4	5	4	4	4
14	B 3078	1	5	4	5	4	0(4)	3
17	B 3033	1	0(3)	2	0	2	0(3)	3
18	B 3031	1	0(4)	4	5	4	0(3)	3
20	B 3076	1	0(3)	3	0	3	0(5)	4
22	B 3058	1	5	5	5	5	4	4
23	B 3079	1	2	2	1	1	1	1
27	B 3047	1	3	2	0(4)	3	0(4)	3
29	C 0008	1	0(3)	3	3	3	0(4)	3

Table 2: ROC data for three mammographers for Group 1, i.e. with Enhancement enabled.

Group2 (without Enhancement)			Mammographer 1		Mammographer 2		Mammographer 3	
Case #	Database	DB Total # of Lesions	BI RAD	LOC	BI RAD	LOC	BI RAD	LOC
3	A 0015	0	2	2	1	1	1	1
4	A 0034	0	2	2	0(3)	2	0(3)	3
5	A 0112	0	2	1	1	1	0(4)	3
8	A 0020	0	2	2	1	1	2	2
9	A 0003	0	3	2	1	1	1	1
13	A 0030	0	2	2	1	1	0(3)	2
15	A 0009	0	2	2	1	1	2	2
16	A 0037	0	2	2	1	1	1	2
17	A 0099	0	0(3)	2	1	1	2	1
18	A 0116	0	0(3)	3	1	1	1	1
21	A 0035	0	0(3)	2	0(4)	3	0(3)	3
23	A 0018	0	2	2	1	1	1	1
24	A 0022	0	2	2	1	1	0(3)	3
27	A 0005	0	0(3)	2	0(3)	2	1	2
30	A 0016	0	2	2	1	1	1	2
1	B 3003	1	1	2	1	1	5	5
2	B 3389	1	2	2	1	1	1	1
6	B 3009	1	0(4)	4	0(3)	2	0(4)	3
7	C 0309	1	4	4	1	1	0(4)	3
10	C 0142	1	0(3)	3	0(3)	2	1	2
11	B 3016	1	0(4)	4	0(3)	2	4	4
12	B 3382	1	2	2	1	1	3	2
14	B 3134	1	5	4	4	4	5	5
19	B 3005	3	0(3)	3	3	3	0(4)	4
20	C 0127	1	0(3)	3	0(4)	3	0(4)	4
22	C 0015	1	0(4)	4	0(4)	4	5	5
25	B 3007	1	3	3	4	3	4	4
26	B 3012	1	5	5	5	5	0(4)	3
28	B 3380	1	0(4)	4	4	4	0(4)	4
29	C 0358	1	5	5	5	4	0(4)	4

Table 3: ROC data for three mammographers for Group 2, i.e. without enhancement.

4.4. ROC Analysis: General Principles

The most widely used method to objectively evaluate the performance of a diagnostic system or the difference in performance between two diagnostic systems is ROC analysis. It compares radiologists' image-based diagnoses with known states of disease and health. In ROC analysis, performance of a diagnostic system is described by the indices of "sensitivity" and "specificity", where "sensitivity" can be expressed as the true-positive fraction (TPF) and "specificity" by the true-negative fraction (TNF) of a diagnosis [16]. In a complimentary way, the false-negative fraction (FNF) and the false-positive fraction (FPF) can be defined as $FNF = 1 - TPF$ and $FPF = 1 - TNF$, respectively, with a similar interpretation. Due to this dependence, it is only necessary to measure one pair of indices, and frequently TPF and FPF are used (as in our study).

The underlying model for ROC analysis is the use of probability density distributions of a radiologist's confidence in a positive diagnosis for a particular diagnostic task for true positive and true negative patients [16]. It is currently accepted that based on a confidence threshold, i.e. a particular level of confidence (LOC) in a positive diagnosis, a diagnosis is considered to be positive, if it exceeds this threshold, and a diagnosis is considered to be negative, if it falls below the threshold. TPF and FPF are then calculated from the probability density distributions as areas under the curves delimited by the confidence threshold (see Figure 10). If the confidence threshold is varied continuously, an ROC curve can be generated from the pair values for TPF and FPF. ROC curves that indicate better decision performance are positioned higher in the unit square spanned by FPF and TPF (higher TPF values for the same FPF values). The area under the ROC curve, A_z , provides a useful summary index for the inherent discrimination performance of a diagnostic system. Thus, A_z is the average value of sensitivity of a corresponding ROC curve, if the specificity of the system is selected randomly between 0.0 and 1.0. Conversely, it can be considered as the average value specificity of a corresponding ROC curve, if the sensitivity of the system is selected randomly between 0.0 and 1.0 [16].

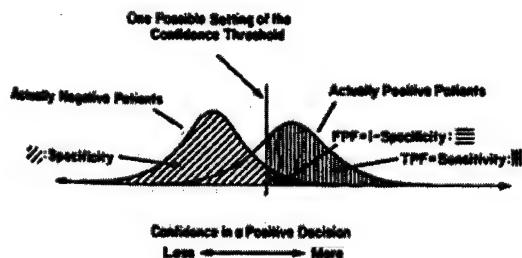


Figure 10: Schematic example of the model that underlies ROC analysis. The bell-shaped curves represent probability density distributions of a radiologist's confidence in a positive diagnosis. A confidence threshold, represented by a vertical line, separates "positive" decisions from "negative" decisions (This figure was reprinted from [16]).

In practice, data for an ROC analysis is obtained by providing a set of rating categories to the radiologist. For a rating scale we chose discrete values from 1 to 5 for the level of confidence (LOC) in a positive diagnosis. The meaning of these values was as follows: (1) definitely or almost definitely negative, (2) probably negative, (3) possibly positive, (4) probably positive, and (5) definitely or almost definitely positive. With this choice the value for the LOC is similar to the standard BI_RAD rating scale used in screening.

To generate an ROC curve from discrete data requires assumptions about the functional form of the curve. The "binormal" model has been widely used in medical imaging. This model includes two adjustable parameters, and it is assumed that each conventional ROC curve has the same functional form as that implied by two "normal" (i.e., Gaussian) decision variable distributions with generally different means and standard deviations [37], [38].

The two adjustable parameters of the binormal ROC curve can be taken to be the y-intercept and the slope of the straight line that represents the ROC curve, when it is plotted on normal-deviate axes. These two parameters, denoted as "a" and "b", can be interpreted as an effective pair of underlying Gaussian distributions as the distance between the means of the two distributions and the standard deviation of the actually negative distribution, respectively with both expressed in units of the standard deviation of the actually positive distribution [16]. With the binormal model, a maximum-likelihood parameter estimation scheme is then used to generate an ROC curve that best represents the data.

If two different diagnostic systems are to be evaluated, the statistical difference of an apparent difference between measured ROC curves is of interest. Testing differences between ROC curves is well described in the literature [39], [40].

4.5. Results from ROC Analysis

In our study, ROC analysis was possible, since the “ground truth” for each case was provided by the mammography database. In general, any enhancement protocol should increase sensitivity, i.e. fraction of true-positives (TPF), without decreasing specificity, i.e. essentially without increasing the fraction of false-positives (FPF) [41]. An initial analysis of the data counted the number of false-positives and true-positives in each group of cases. Before a lesion was considered being diagnosed as malignant or benign, the LOC value was thresholded [16]. The threshold value influences the shape of the ROC curve and its interpretation. For example, if the threshold for the level of confidence was chosen to be 3, meaning that lesions with a LOC greater or equal 3 were considered as malignant, then the average TPF was found to be 0.667 with enhancement, and TPF = 0.569 without enhancement. This observed increase in sensitivity is encouraging, though it was accompanied by a slight increase in the fraction of false-positives (0.222 compared to 0.178). The latter is not too surprising, *since the applied enhancement protocol only used dyadic spline wavelets* with the non-linear sigmoidal enhancement function, which is certainly not optimal for all types of lesions. We believe that dyadic spline wavelet expansions are best used to enhance microcalcifications. If the analysis of the data only focused on microcalcifications, then we observed TPF = 0.417 with enhancement compared to TPF = 0.222 without enhancement. No increase or decrease in FPF was noticed! The last finding supports the promise for future research to design specific enhancement protocols for each mammographic feature. Table 4 summarizes initial results of the ROC study using the single basis function described earlier in Section 2.3.

With Enhancement (all types of lesions)		Without Enhancement (all types of lesions)	
TPF	FPF	TPF	FPF
0.667	0.233	0.569	0.178
With Enhancement (Micros only)		Without Enhancement (Micros only)	
TPF	FPF	TPF	FPF
0.417	0.0	0.222	0.0

Table 4: Results of preliminary ROC study. TPF refers to the fraction of true-positives and FPF to the fraction of false-positives.

A more thorough analysis of the data was undertaken by using the *ROCKIT* software developed by a research group led by Charles Metz at the University of Chicago [42], [43]. This software package was written to analyze data from ROC studies and to generate corresponding ROC curves. More specifically, the purpose of *ROCKIT* is to calculate maximum-likelihood estimates of the parameters of a conventional “binormal” model for the input data, to calculate maximum-likelihood estimates of the parameters of a “bivariate binormal” model for data from two potentially correlated diagnostic tests and, thus, to estimate the binormal ROC curves implied by those data and their correlation; and to calculate the statistical significance of the difference between two ROC curve estimates using any one of three distinct statistical tests:

1. **The Bivariate Test:** A bivariate Chi-square test of the simultaneous differences between the “*a*” parameters and between the “*b*” parameters of the two ROC curves. (*Null hypothesis:* the data sets arose from the same binormal ROC curve.)
2. **The Area Test:** A univariate z-score test of the difference between the areas under the two ROC curves. (*Null hypothesis:* the data sets arose from binormal ROC curves with equal areas beneath them.)
3. **The TPF Test:** A univariate z-score test of the difference between the true-positive fractions (TPFs) on the two ROC curves at a selected false-positive fraction (FPF). (*Null hypothesis:* the data sets arose from binormal ROC curves having the same TPF at the selected FPF.)

Three types of input data are allowed for statistical testing of the differences between ROC curves:

1. Unpaired (uncorrelated) test results. The two “conditions” are applied to independent case samples — for example, from two different diagnostic tests performed on the different patients, from two different radiologists who make probability judgments concerning the presence of a specified disease in different images, etc.;
2. Fully paired (correlated) test results, in which data from both of two conditions are available for each case in a single case sample. The two “conditions” in each test-result pair could correspond, for example, to two different diagnostic tests performed on the same patient, to two different radiologists who make probability judgments concerning the presence of a specified disease in the same image, etc.; and
3. Partially-paired test results — for example, two different diagnostic tests performed on the same patient sample and on some additional patients who received only one of the diagnostic tests.

ROCKIT assumes that the population ROC curve for each condition plots as a straight line on “normal-deviate” axes, or equivalently, that the input data follow normal distributions after some unknown monotonic transformation [16]. ROC curves measured in a broad variety of fields demonstrate this “binormal” form [44], [45], and [46]. The assumption may be satisfied even when the raw data have multimodal and/or skewed distributions [43], [42].

Using the *ROCKIT* software the analysis was first applied independently to the datasets for Group 1 and Group 2 for each of the three radiologists. Unfortunately, this approach did not allow us to compare the diagnostic performance for the two diagnostic systems (softcopy display with and without enhancement). The reason for that was that the analysis for, at least one group of cases could not be completed, since the data was found to be degenerate [41]. In this case, the result of the ROC analysis would be a straight line with a constant value for TPF, and, therefore the software aborts processing to avoid meaningless output. According to the authors of the software, a degenerate data distribution can be found, if the number of samples is too small or in datasets with many tied values [43].

Since the number of cases could not be increased after conducting the study, and in order to obtain more complete results, we decided to apply the analysis to the union of data from all three radiologists. This was justified by the fact that all three radiologists came from the same population with a similar level of experience. Thus, their performance should be similar under the same conditions, and the data could be treated as independent samples (unpaired data). If the data did not have to be pooled, it would have been unpaired, since the two different conditions were applied to different sample cases. Nevertheless, we are well aware that the statistical significance of the results must be interpreted carefully. For future ROC studies we plan to increase the number of cases, in order to avoid such a problem. To check on our assumption of independent samples (unpaired data) and for completion we also repeated the analysis with the input as paired data. These results are included in this chapter as well.

For the analysis Group 1 (with enhancement) was set as Condition 1 and Group 2 (without enhancement) was considered as Condition 2. The resulting ROC curves for data analyzed as unpaired are shown in Figure 11. Their corresponding values for FPF and TPF are given in Table 5. Finally, the most important results of ROC analysis, the binormal parameters a , b , and the area under the ROC curve A_z with their corresponding standard errors, 95% confidence intervals, and correlation of a and b are summarized for unpaired data in Table 6. Note that the 95% confidence intervals are symmetric for the binormal parameters a and b , but asymmetric for the area index A_z . The corresponding results from the analysis as paired data follow directly afterwards. ROC curves are shown in Figure 12, FPF and TPF values in Table 7, and parameters a , b , and A_z together with their corresponding standard errors, 95% confidence intervals, and correlation of a and b in Table 8.

ROC Curves for Data with and without Enhancement

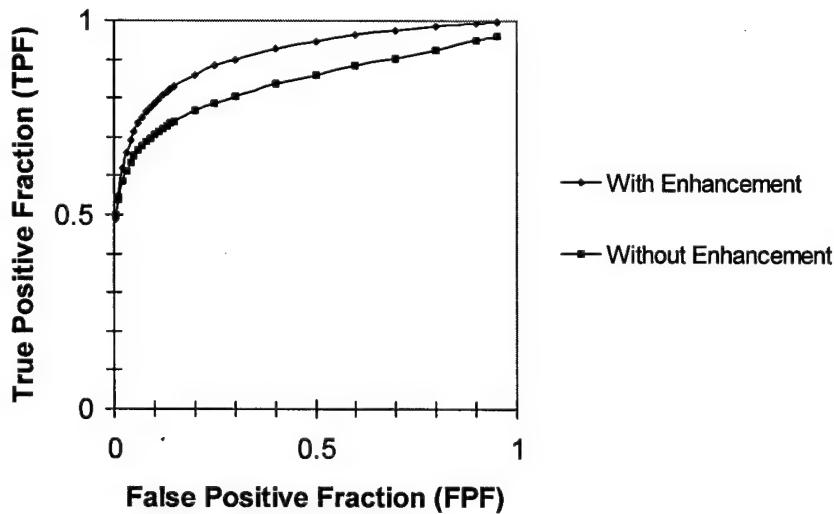


Figure 11: ROC curves for data with Condition 1 (with enhancement) and Condition 2 (without enhancement) analyzed as unpaired data (independent analysis).

FPF	TPF 1	TPF 2	FPF	TPF 1	TPF 2
0.005	0.4886	0.4989	0.13	0.8155	0.7282
0.01	0.5521	0.5407	0.14	0.8232	0.7346
0.02	0.6199	0.5859	0.15	0.8304	0.7406
0.03	0.6612	0.614	0.2	0.86	0.7665
0.04	0.6911	0.6347	0.25	0.8825	0.7874
0.05	0.7145	0.6514	0.3	0.9003	0.8053
0.06	0.7338	0.6653	0.4	0.9274	0.8352
0.07	0.7501	0.6773	0.5	0.9472	0.8602
0.08	0.7642	0.6879	0.6	0.9625	0.8825
0.09	0.7767	0.6974	0.7	0.9746	0.9035
0.1	0.7878	0.7061	0.8	0.9845	0.9244
0.11	0.7979	0.714	0.9	0.9926	0.9475
0.12	0.8071	0.7213	0.95	0.9962	0.9619

Table 5: Values for false-positive fractions (FPF) and true-positive fractions (TPF) for Condition 1 (with enhancement, TPF 1) and Condition 2 (without enhancement, TPF 2) analyzed as unpaired data (independent analysis).

Condition 1 (With Enhancement)			Condition 2 (Without Enhancement)		
Binormal Parameter a	Binormal Parameter b	Area under ROC Curve A_z	Binormal Parameter a	Binormal Parameter b	Area under ROC Curve A_z
1.6183	0.6393	0.9136	1.0813	0.4208	0.8405
Standard Error a	Standard Error b	Standard Error A_z	Standard Error a	Standard Error b	Standard Error A_z
0.3162	0.2093	0.0325	0.2329	0.1307	0.0475
95% Confidence Interval for a	95% Confidence Interval for b	95% Confidence Interval for A_z	95% Confidence Interval for a	95% Confidence Interval for b	95% Confidence Interval for A_z
(0.9986, 2.2381)	(0.2291, 1.0495)	(0.8312, 0.9615)	(0.6247, 1.5379)	(0.1647, 0.6770)	(0.7301, 0.9162)
Correlation(a, b)			Correlation(a, b)		
0.6544			0.4989		

Table 6: Binormal parameters a , b , area under ROC curve A_z with their corresponding standard errors, 95% confidence intervals, and correlation(a , b) for Condition 1 (with enhancement) and Condition 2 (without enhancement) analyzed as unpaired data (independent analysis).

ROC Curves for Data

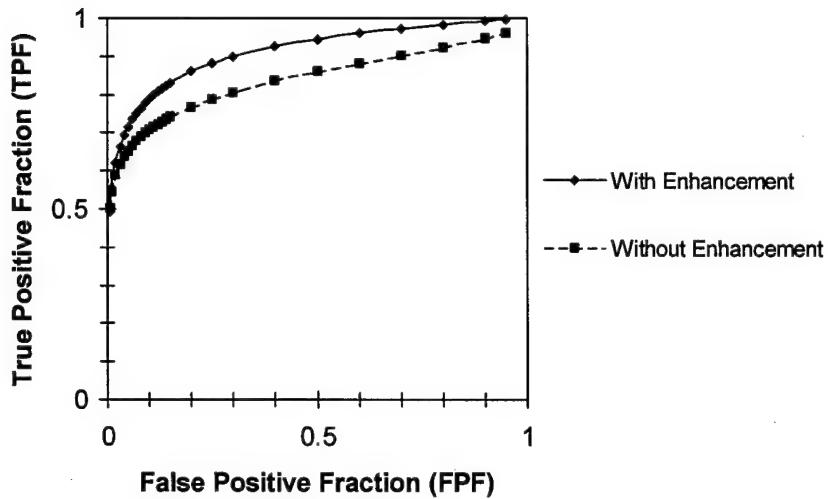


Figure 12: ROC curves for data with Condition 1 (with enhancement) and Condition 2 (without enhancement) analyzed as paired data (correlated analysis).

FPF	TPF 1	TPF 2	FPF	TPF 1	TPF 2
0.005	0.494	0.5036	0.13	0.8155	0.7304
0.01	0.5565	0.5451	0.14	0.8232	0.7367
0.02	0.6232	0.5898	0.15	0.8303	0.7426
0.03	0.6638	0.6176	0.2	0.8595	0.7682
0.04	0.6932	0.6381	0.25	0.8817	0.7889
0.05	0.7162	0.6545	0.3	0.8994	0.8066
0.06	0.7351	0.6683	0.4	0.9263	0.8361
0.07	0.7512	0.6801	0.5	0.9461	0.8608
0.08	0.7651	0.6906	0.6	0.9614	0.8829
0.09	0.7774	0.7	0.7	0.9737	0.9036
0.1	0.7883	0.7086	0.8	0.9838	0.9244
0.11	0.7982	0.7164	0.9	0.9922	0.9472
0.12	0.8073	0.7236	0.95	0.9959	0.9617

Table 7: Values for false-positive fractions (FPF) and true-positive fractions (TPF) for Condition 1 (with enhancement, TPF 1) and Condition 2 (without enhancement, TPF 2) analyzed as paired data (correlated analysis).

Condition 1 (With Enhancement)			Condition 2 (Without Enhancement)		
Binormal Parameter a	Binormal Parameter b	Area under ROC Curve A_z	Binormal Parameter a	Binormal Parameter b	Area under ROC Curve A_z
1.6084	0.6302	0.9132	1.0839	0.4172	0.8414
Standard Error a	Standard Error b	Standard Error A_z	Standard Error a	Standard Error b	Standard Error A_z
0.3137	0.2072	0.0327	0.233	0.1302	0.0474
95% Confidence Interval for a	95% Confidence Interval for b	95% Confidence Interval for A_z	95% Confidence Interval for a	95% Confidence Interval for b	95% Confidence Interval for A_z
(0.9936, 2.2232)	(0.2240, 1.0363)	(0.8304, 0.9613)	(0.6272, 1.5407)	(0.1620, 0.6724)	(0.7311, 0.9169)
Correlation(a, b)			Correlation(a, b)		
0.6506			0.4995		

Correlation of A_z for Condition 1 and A_z for Condition 2: -0.0922

Table 8: Binormal parameters a , b , area under ROC curve A_z with their corresponding standard errors, 95% confidence intervals, and correlation(a , b) for Condition 1 (with enhancement) and Condition 2 (without enhancement) analyzed as paired data (correlated analysis).

4.6. Discussion

As seen from the analysis for unpaired data, the value for the area under the ROC curve A_z was by 8.7% larger for Condition 1 (with enhancement) than it was for Condition 2 (without enhancement). In all cases the standard error for A_z was between 0.03 and 0.05, which was rather small. Though the 95% confidence intervals for A_z overlapped, there was a clear tendency that diagnostic performance improved *with enhancement* in comparison with diagnosis *without enhancement*. All ROC curves lay high in the unit square of FPF and TPF, which corresponded to accurate diagnostic performances in general, but the curve for Condition 1 was positioned slightly higher (see Figure 11).

Similar results were generally obtained for the analysis as paired data. The increase in A_z for Condition 1 with respect to Condition 2 was 8.5 %, but there was an overlap of the 95% confidence intervals for A_z as well. The ROC curve for Condition 1 was also positioned slightly higher than the one for Condition 2 (see Figure 12). Values for a , b , and A_z were very similar for both types of analysis. Hence, the same tendency of improved diagnostic performance *with enhancement* compared to diagnosis *without enhancement* can be inferred.

The observed increase of the summary index A_z within statistical errors and the higher position of the ROC curve for diagnosis with enhancement encourage us to further pursue the application of enhancement protocols for mammographic screening. We are aware of the fact that there always are inherent sources of variability in the index A_z , such as a “case-sample” component due to random variations in the difficulty of the cases included in an ROC experiment, a “between-reader” component due to random variations in the skills of the observers participating in the experiment, and a “within-reader” component associated with each reader’s inability to reproduce her/his diagnosis of every case on repeated readings [16]. In addition, we were not able to analyze the data for each radiologist separately due to data degeneracy as mentioned above. The latter has diminished the statistical significance of our results obtained from the analysis of all data combined, since not all samples were completely independent.

Hence, for future ROC studies we plan to increase the number of cases to avoid degenerate datasets for the analysis and to increase the statistical power of the experiment.

Aside from statistical considerations and the cautious interpretation of the results of this study we know that our prototype test bed software tool can be further optimized. To improve multi-scale contrast enhancement the idea is to develop feature specific enhancement protocols with different bases and associated non-linear functions for each distinct mammographic feature, such as microcalcifications, masses, and spicular lesions. The enhancement protocol used for this experiment, dyadic spline wavelets with non-linear sigmoidal function, was suggested to work best for microcalcifications according to our previous work with multi-scale expansions [2], [25]. The results of this first ROC experiment seem to confirm our expectations.

5. CONCLUSIONS AND FUTURE WORK

We have reported on the successful completion of the first receiver operating characteristics (ROC) study to evaluate the benefits of contrast enhancement via overcomplete multi-scale expansions of mammograms. The study was carried out in collaboration with radiologists at the Breast Imaging Center in Columbia-Presbyterian Medical Center and the Biomedical Imaging Laboratory of Columbia University.

In continuation of our previous work in digital mammography, an enhancement protocol using a dyadic spline wavelet as the basis for multi-scale expansion and an associated non-linear sigmoidal enhancement function was designed. Suspicious areas (ROIs) of digitized mammograms were decomposed onto a multi-scale basis to obtain coefficients at distinct subbands. Coefficients were modified by applying a non-linear sigmoidal function. Two parameters could be adjusted to change the nature of enhancement. Image reconstruction from modified coefficients occurred in nearly real time through an interactive interface running on a high-resolution digital mammography workstation. To visualize raw data of digitized mammograms at the highest possible contrast and spatial resolutions, 16-Bit BARCO/Metheus framebuffers together with a dual headed high-resolution MegaScan grayscale monitor were utilized in hardware. We incorporated specialized software function calls to directly access the video framebuffer for fast/smooth image display and update.

To quantify the performance of our multi-scale based processing technique in terms of overall sensitivity and specificity, an ROC study was designed and conducted with three radiologists from Columbia-Presbyterian Medical Center specialized in mammography. Conventional ROC curves were generated and significant statistical parameters determined. The area under the ROC curve A_z was used as a summary index to quantify overall specificity and sensitivity of the two diagnostic systems [16]. Unfortunately, it was not possible to analyze datasets for each of three mammographers separately due to data degeneracy. Nevertheless, analyzing all the data together yielded a slight increase (8.7%) in the area A_z for diagnosis with enhancement compared to diagnosis without. Despite the limited statistical significance of this result, it encourages us to further investigate the application of multi-scale methods for contrast enhancement of mammograms. More extensive ROC studies with a larger number of cases are planned to further evaluate the benefits of such processing techniques.

Ancillary to statistical results, we received very positive feedback from the participating radiologists, who expressed great interest in using the interactive display tool and acknowledged a marked improvement in image quality, when enhancement was applied.

The current enhancement protocol works best for the detection/enhancement of microcalcifications. Future directions of work include the expansion of the choice of enhancement protocols to a menu of feature specific enhancement algorithms tailored for each mammographic feature, such as microcalcifications, masses, and spicular lesions, e.g. the application of brushlet functions [47], [48] to mammograms with spicular lesions. In addition, the investigation of a range of optimal enhancement parameters and the optimization of our interface software tool comprise further projects. Our "dream" is to present a clinical interface, where specific enhancement protocols can be selected by a physician by only "pushing a button on the screen". We envision that through such a clinical interface the diagnostic performance of radiologists in screening digital mammograms could be substantially improved, both in terms of cost and quality.

6. ACKNOWLEDGEMENT

This work was supported by the Breast Cancer Research Program of the Department of Defense U.S. Army Medical Research and Material Command, Award Number DAMD17-93-J-3003 and the Whitaker Foundation.

7. REFERENCES

- [1] I. Koren and A. Laine, "A discrete dyadic wavelet transform for multidimensional feature analysis," in *Time Frequency and Wavelets in Biomedical Signal Processing, IEEE Press series in biomedical engineering*, M. Akay, Ed. Piscataway, NJ: IEEE Press, 1998, pp. 425-448.
- [2] A. F. Laine, S. Schuler, J. Fan, and W. Huda, "Mammographic feature enhancement by multiscale analysis," *IEEE Transactions on Medical Imaging*, vol. 13, pp. 725-740, 1994.
- [3] J. Fan and A. F. Laine, "Multiscale contrast enhancement and denoising in digital radiographs," in *Wavelets in Medicine and Biology*, A. Aldroubi and M. Unser, Eds. Boca Raton, FL: CRC Press, 1996, pp. 163-189.
- [4] D. Brzakovic, X. M. Luo, and P. Brzakovic, "An approach to automated detection of tumors in mammograms," *IEEE Transactions on Medical Imaging*, vol. 9, pp. 233-241, 1991.
- [5] H. Yoshida, K. Doi, and R. M. Nishikawa, "Automated detection of clustered microcalcifications in digital mammograms using wavelet transform techniques," in *Proceedings of the SPIE*, vol. 2167, pp. 868-886, 1994.
- [6] H. Yoshida, W. Zhang, W. Cai, K. Doi, R. M. Nishikawa, and M. L. Giger, "Optimizing wavelet transform based on supervised learning for detection of microcalcifications in digital mammograms," in *Proceedings of the IEEE International Conference on Image Processing*, vol. 3, Washington, D.C., pp. 152-155, 1995.

- [7] D. Wei, H.-P. Chan, M. A. Helvie, B. Sahiner, and N, "Classification of mass and normal breast tissue on digital mammograms: multiresolution texture analysis," *Medical Physics*, vol. 22, pp. 1501-1513, 1995.
- [8] L. Li, W. Qian, and L. P. Clark, "X-ray medical image processing using directional wavelet transform," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, Atlanta, GA, pp. 2251-2254, 1996.
- [9] A. F. Laine, W. Huda, D. Chen, and J. Harris, "Segmentation of masses using continuous scale representations," in *Proceedings of the Third International Workshop on Mammography*, Chicago, IL, pp. 447-450, 1996.
- [10] R. N. Strickland and H. I. Hahn, "Wavelet transforms for detecting microcalcifications in mammograms," *IEEE Transactions on Medical Imaging*, vol. 15, pp. 218-229, 1996.
- [11] W. Qian, L. P. Clarke, M. Kallergi, and H.-D. Li, "Tree-structured nonlinear filter and wavelet transform for microcalcification segmentation in mammography," in *Proceedings of the SPIE: Biomedical Image Processing and Biomedical Visualization*, vol. 1905, San Jose, CA, pp. 520-590, 1995.
- [12] Y. Xing, W. Huda, A. F. Laine, and J. Fan, "Simulated phantom images for optimizing wavelet-based image processing algorithms in mammography," in *Proceedings of the SPIE: Mathematical Methods in Medical Imaging III*, vol. 2299, San Diego, CA, pp. 207-217, 1994.
- [13] Y. Xing, W. Huda, A. Laine, J. Fan, and B. Steinbach, "Comparison of a dyadic wavelet image enhancement algorithm with unsharp masking and median filtering," in *Proceedings of the SPIE: Medical Imaging - Image Processing*, vol. 2434, San Diego, CA, pp. 718-729, 1995.
- [14] D. Chen, C.-M. Chang, and A. Laine, "Detection and enhancement of small masses via precision multiscale analysis," in *Proceedings of the Third Asian Conference on Computer Vision: Computer Vision - ACCV '98*, vol. 1, Hong Kong, PRC, pp. 192-199, 1998.
- [15] N. Karssemeijer and G. M. teBrake, "Detection of stellate distortions in mammograms," *IEEE Transactions on Medical Imaging*, vol. 15, pp. 611-619, 1996.
- [16] C. E. Metz, "ROC methodology in radiologic imaging," *Investigative Radiology*, vol. 21, pp. 720-733, 1986.
- [17] B. J. Betts, J. Li, A. Aiyer, S. M. Perlmutter, P. C. Cosman, R. M. Gray, R. A. Olshen, and al., "Image Quality," Stanford University, Stanford, Final Report For U.S. Army Medical Research and Material Command, Fort Detrick, Maryland, November 18 pp. 1-71, 1998.
- [18] H. Nab, N. Karssemeijer, L. Van Erning, and J. Hendriks, "Comparison of digital and conventional mammography: a ROC study of 270 mammograms," *Medical Informatics*, vol. 17, pp. 125-131, 1992.
- [19] W. Qian, L. Li, and L. P. Clarke, "Image feature extraction for mass detection in digital mammography: Influence of wavelet analysis," *Medical Physics*, vol. 26, pp. 402-408, 1999.
- [20] H.-P. Chan, B. Sahiner, R. Wagner, and N. Petrick, "Effects of sample size on classifier design for computer-aided diagnosis," in *Proceedings of the SPIE: Medical Imaging - Image Processing*, vol. 3338, pt.1-2, San Diego, CA, USA, pp. 845-858, 1998.
- [21] M. Unser and A. Aldroubi, "A review of wavelets in biomedical applications," *Proceedings of the IEEE*, vol. 84, pp. 626-638, 1996.
- [22] M. Holschneider and R. Kronland-Martinet, "A real-time algorithm for signal analysis with the help of the wavelet transform," in *Wavelets: Time-frequency Methods and Phase Space*, Berlin, Germany, pp. 286-304, 1990.
- [23] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multiscale transforms," *IEEE Transactions on Information Theory*, vol. 38, pp. 587-607, 1992.
- [24] S. D. Marco and J. Weiss, "M-band wavepacket-based transient signal detector using a translation-invariant wavelet," *Optical Engineering*, vol. 33, pp. 2175-2182, 1994.
- [25] A. F. Laine, J. Fan, and W. Yang, "Wavelets for contrast enhancement of digital mammography," *IEEE Engineering in Medicine and Biology Society Magazine*, vol. 14, pp. 536-550, 1995.
- [26] A. F. Laine, J. Fan, and S. Schuler, "A framework for contrast enhancement by dyadic wavelet analysis," in *Digital Mammography*, A. G. Gale, S. M. Astley, D. R. Dance, and A. Y. Cairns, Eds. Amsterdam, The Netherlands: Elsevier, 1994, pp. 91-100.
- [27] C.-M. Chang and A. F. Laine, "Enhancement of mammograms from oriented information," in *Proceedings of the IEEE International Conference on Image Processing*, vol. 3, Santa Barbara, CA, pp. 524-527, 1997.
- [28] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, CA: Academic Press, 1998.
- [29] P. G. Tahoces, J. Correa, M. Souto, and C. G. and, "Enhancement of chest and breast radiographs by automatic spatial filtering," *IEEE Transactions on Medical Imaging*, vol. 10, pp. 330-335, 1991.
- [30] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: Siam, 1992.
- [31] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 710-732, 1992.

- [32] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Transactions on Information Theory*, vol. 38, pp. 617-643, 1992.
- [33] I. Koren, A. Laine, F. Taylor, and M. Lewis, "Interactive wavelet processing and techniques applied to digital mammography," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, Atlanta, GA, pp. 1415-1418, 1996.
- [34] W. B. Richardson Jr., "Nonlinear filtering and multiscale texture discrimination for mammograms," in *Proceedings of the SPIE: Mathematical Methods in Medical Imaging*, vol. 1768, San Diego, CA, pp. 293-305, 1992.
- [35] D. L. Donoho and I. M. Johnstone, "Threshold selection for wavelet shrinkage of noisy data," in *Proc. 16th Annual Int. Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 1, pp. A24-25, 1994.
- [36] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, pp. 425-455, 1994.
- [37] D. M. Green and J. A. Swets, *Signal detection theory and psychophysics*. New York: Wiley, 1966.
- [38] J. A. Swets, "ROC analysis applied to the evaluation of medical imaging techniques," *Investigative Radiology*, vol. 14, pp. 109-21, 1979.
- [39] B. J. McNeil and J. A. Hanley, "Statistical approaches to the analysis of receiver operating characteristic (ROC) curves," *Medical Decision Making*, vol. 4, pp. 137-150, 1984.
- [40] J. A. Swets and R. M. Pickett, *Evaluation of diagnostic systems: Methods from signal detection theory*. New York, NY: Academic Press, 1982.
- [41] C. E. Metz, "Some practical issues of experimental design and data analysis in radiological ROC studies," *Investigative Radiology*, vol. 24, pp. 234-245, 1989.
- [42] C. E. Metz, "ROCKIT 0.9B", Beta Version, Department of Radiology, University of Chicago, Chicago, IL, 1998.
- [43] C. E. Metz, B. A. Herman, and C. A. Roe, "Statistical comparison of two ROC curve estimates obtained from partially-paired datasets," *Medical Decision Making*, vol. 18, pp. 110-121, 1998.
- [44] J. A. Swets, "Form of empirical ROCs in discrimination and diagnostic tasks: implications for theory and measurement of performance," *Psych Bull*, vol. 99, pp. 181-198, 1986.
- [45] J. A. Hanley, "The robustness of the "binormal" assumptions used in fitting ROC curves," *Medical Decision Making*, vol. 8, pp. 197-203, 1988.
- [46] K. O. Hajian-Tilaki, J. A. Hanley, L. Joseph, and J.-P. Collet, "A comparison of parametric and nonparametric approaches to ROC analysis of quantitative diagnostic tests," *Medical Decision Making*, vol. 17, pp. 94-107, 1997.
- [47] F. Meyer and R. R. Coifman, "Brushlets: A tool for directional image analysis and image compression," *Applied and computational harmonic analysis*, vol. 4, pp. 147-187, 1997.
- [48] E. Angelini, A. Laine, S. Takuma, and S. Homma, "Directional representations of 4D echocardiography for temporal quantification of LV volumes," in *Medical Imaging and Computer-Assisted Intervention - MICCAI'99*, Cambridge, England, pp. 430-440, 1999.

Visualization of mammograms via fusion of enhanced features

Iztok Koren^a, Andrew Laine^b, Suzanne Smith^c, Edward Nickoloff^c, and Fred Taylor^a

^aDepartment of Electrical and Computer Engineering,
University of Florida, Gainesville, FL 32611, USA

^bCenter for Biomedical Engineering,
Columbia University, New York, NY 10027, USA

^cDepartment of Radiology,
Columbia-Presbyterian Medical Center, New York, NY 10032, USA

Image enhancement in mammography is typically concerned with either general visibility of all features or conspicuity of a specific sign of malignancy. We describe a synthesis of the two approaches through fusion of locally enhanced microcalcifications, circumscribed masses, and stellate lesions. Both local processing and image fusion are performed within a single wavelet transform framework which contributes to the computational efficiency of the method. The algorithm not only allows for efficient combination of specific features of importance, but also provides a flexible framework for incorporation of distinct enhancement methods and their independent optimization.

Mammography, contrast enhancement, image fusion, wavelet transform.

1. INTRODUCTION

In general, mammographic image enhancement methods target either visualization of all features in an image [1, 2, 3, 4] or visibility of specific features of importance such as microcalcifications [5].

Methods from the first category are not optimized for a specific type of cancer, and are often developed for a framework more general than mammography alone. The second category approaches can be quite successful in their area of specialization; however, in order to process mammograms for presence of distinct features, independent application of different algorithms could result in both larger number of images to be interpreted by a radiologist and increased computational complexity.

Here, we present an approach which overcomes these shortcomings and problematic limitations via synthesis of the two paradigms by means of image fusion. The algorithm consists of two major steps: (1) wavelet coefficients are modified distinctly for each type of malignancy; (2) the obtained multiple sets of wavelet coefficients are fused into a single

This work was supported by the U.S. Army Medical Research and Materiel Command under DAMD17-96-1-6093 and DAMD17-93-J-3003, and by the Whitaker Foundation.

set from which a reconstruction is computed. The scheme allows efficient deployment of an enhancement strategy appropriate for clinical screening protocols: an enhancement algorithm is first developed for each specific type of feature independently, and the results are then combined using an appropriate fusion strategy.

2. WAVELET TRANSFORM

Wavelet based methods are particularly well suited for processing of mammograms since mammographic features greatly vary in shape and size. Commonly used orthogonal and biorthogonal wavelet transforms, however, may not be the best tool for mammographic image enhancement because their lack of translation invariance can lead to artifacts possibly affecting a radiologist's interpretation. Translation-invariant but overcomplete wavelet representations avoid artifacts and have been successfully used for processing of mammograms [1, 2, 5].

Rotation invariance is another desirable property of wavelet decompositions. The concept of steerability [6] has been utilized for construction of wavelet transforms enabling rotation-invariant processing of mammograms [3]. Our scheme is built around a multiscale spline derivative-based transform which, in addition to being translation-invariant and approximately steerable, is also suitable for non-linear methods of enhancement.

We use x - y separable wavelets

$$\psi(x, y) = \frac{d^d \beta_{p+d}(x)}{dx^d} \beta_{p+d}(y), \quad (1)$$

where $\beta_p(x)$ denotes a central B-spline of order p , and limit ourselves to first and second derivatives $d \in \{1, 2\}$. Figure 1 shows wavelets with $p=3$.

A rotation of wavelet $\psi(x, y)$ by angle θ can be expressed as

$$\psi^\theta(x, y) \simeq \sum_{i=0}^d \binom{d}{i} n_x^{d-i} n_y^i \frac{d^{d-i} \beta_{p+d}(x)}{dx^{d-i}} \frac{d^i \beta_{p+d}(y)}{dy^i},$$

where $\vec{n} = (\cos \theta, \sin \theta) = (n_x, n_y)$. The terms $\frac{d^{d-i} \beta_{p+d}(x)}{dx^{d-i}} \frac{d^i \beta_{p+d}(y)}{dy^i}$ represent basis functions needed to approximately steer wavelet $\psi(x, y)$. A dyadic wavelet transform using these basis functions can be implemented as a filter bank consisting of one-dimensional filters only [7].

3. ENHANCEMENT OF MAMMOGRAPHIC FEATURES

3.1. Microcalcifications

Microcalcifications appear on mammograms in approximately half of breast cancer cases. The assessment of shape, number, and distribution of microcalcifications is important for a radiologist to reach diagnosis. Microcalcifications are smaller than 1 mm in size and can be difficult to locate when they are superimposed on dense breast tissue.

Several techniques have been developed to improve the visibility of microcalcifications [5, 8, 9]. The approach devised by Strickland and Hahn [5] is particularly well suited

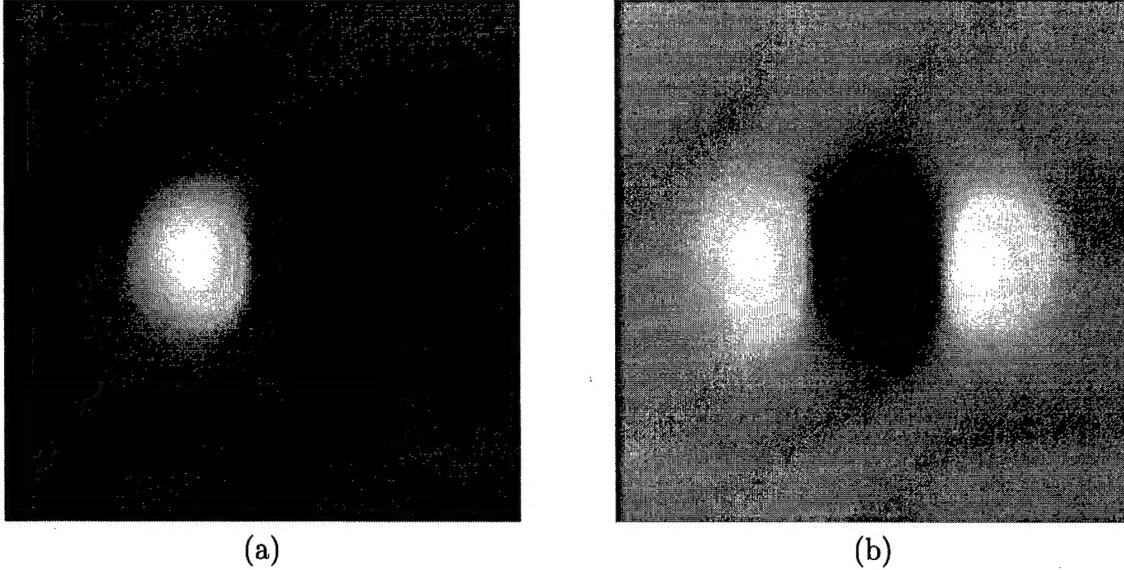


Figure 1. Spline derivatives in the x -axis direction. (a) Wavelet equal to the first derivative of a quartic spline. (b) Wavelet equal to the second derivative of a quintic spline.

for our framework: they used an undecimated wavelet transform to approximate second derivatives of a Gaussian probability density function for a multiscale matched filtering for presence of microcalcifications.

Strickland and Hahn based their method on the observation that the average microcalcification can be modeled by a circularly symmetric Gaussian function. We take advantage of this fact to model microcalcifications by central B-splines. Using B-spline approximations of a Gaussian function, the assumption that a Gaussian object is visible approximately over $\pm\sigma$ pixels [5], and the fact that mammograms in the University of Florida database were digitized at $116\mu\text{m}$ resolution, four levels of the transform described in Section 2 with, for example, $p=3$ are needed to encompass different sizes of microcalcifications. The wavelet decomposition including voices at scales 3 and 6 (corresponding to Strickland and Hahn's octaves "2.5" and "3.5") was obtained from relations between central B-splines at integer scales [10].

The wavelet decomposition enables approximations both to the second derivatives of Gaussian along x and y directions and to Laplacian of Gaussian across distinct scales employed by Strickland and Hahn. We proceed in a similar fashion: the two outputs per scale are thresholded independently, all binary results are then combined, a circular region centered at detected pixel locations is next multiplied by a gain, and, finally, the modified transform coefficients are used for image fusion.

3.2. Circumscribed Masses

Almost half of missed cancers appear on mammograms as masses. Perception is a problem particularly for patients with dense fibroglandular patterns. The detection of masses can be especially difficult because of their small size and subtle contrast compared with normal breast structures.

Fan and Laine [2] developed a discrete dyadic wavelet transform based algorithm suitable for enhancement of masses. They constructed an approximation to Laplacian of Gaussian across dyadic scales for an isotropic input to a piecewise linear enhancement function.

An approximation to a Laplacian of Gaussian across dyadic scales is easy to obtain using multiscale spline derivatives from Section 2: basis functions $\frac{d^{d-i}\beta_{p+d}(x)}{dx^{d-i}} \frac{d^i\beta_{p+d}(y)}{dy^i}$ with $d = 2$ and $i \in \{1, 2\}$ approximate the second derivative of a Gaussian function along directions of x and y axis. The appropriate transform coefficient at each dyadic scales are then added and their sum input to the piecewise linear function [2]

$$C(x) = \begin{cases} x - (K-1)T & \text{if } x < -T \\ Kx & \text{if } |x| \leq T \\ x + (K-1)T & \text{if } x > T \end{cases}$$

used at each level $m+1$ of the transform separately. Due to the expected size of masses, levels greater than 4 are enhanced more aggressively.

The multiplicative factor obtained as the ratio between the output and input of the enhancement function is next applied to the original wavelet coefficients before fusion and the associated inverse wavelet transform are carried out.

3.3. Stellate Lesions

It is important for radiologists to identify stellate lesions since their presence is a serious indicator of malignancy. Stellate lesions vary in size and subtlety and, in addition, do not have a clear boundary, making them difficult to detect.

In the development of our algorithm, we utilized an observation made by Kegelmeyer *et al.* about the distortion of edge orientation distribution induced by a stellate lesion [11]. Normal mammograms show a roughly radial pattern with structure radiating from the nipple to the chest wall. A stellate lesion not only changes this pattern, but also creates another center from which rays radiate.

The wavelet transform from Section 2 allows directional analysis using approximations to both first and second steerable derivatives of a Gaussian. A multiscale derivative-pair quadratic feature detector was computed by finding the maximum of the local oriented energy with respect to angle θ ,

$$E_{2^m}^\theta(x, y) = \sqrt{(W1_{2^m}^\theta s(x, y))^2 + (W2_{2^m}^\theta s(x, y))^2}, \quad (2)$$

where $W1_{2^m}^\theta s(x, y)$ and $W2_{2^m}^\theta s(x, y)$ denote wavelet decompositions using first (Equation (1) with $d = 1$) and second (Equation (1) with $d = 2$) derivative wavelet, respectively, steered to angle θ . The angle that maximizes the local oriented energy (2) represents orientation at pixel location (x, y) .

Similar to the method from Section 3.1, processing is carried out within windows of scale dependent size: 1-norm of differences between the local and average orientations was computed in the window and used as a measure of orientation nonuniformity. Soft thresholding as a function of the orientation nonuniformity measure was next applied to the transform coefficients at each dyadic scale independently. The altered coefficients are then included for fusion and reconstruction.

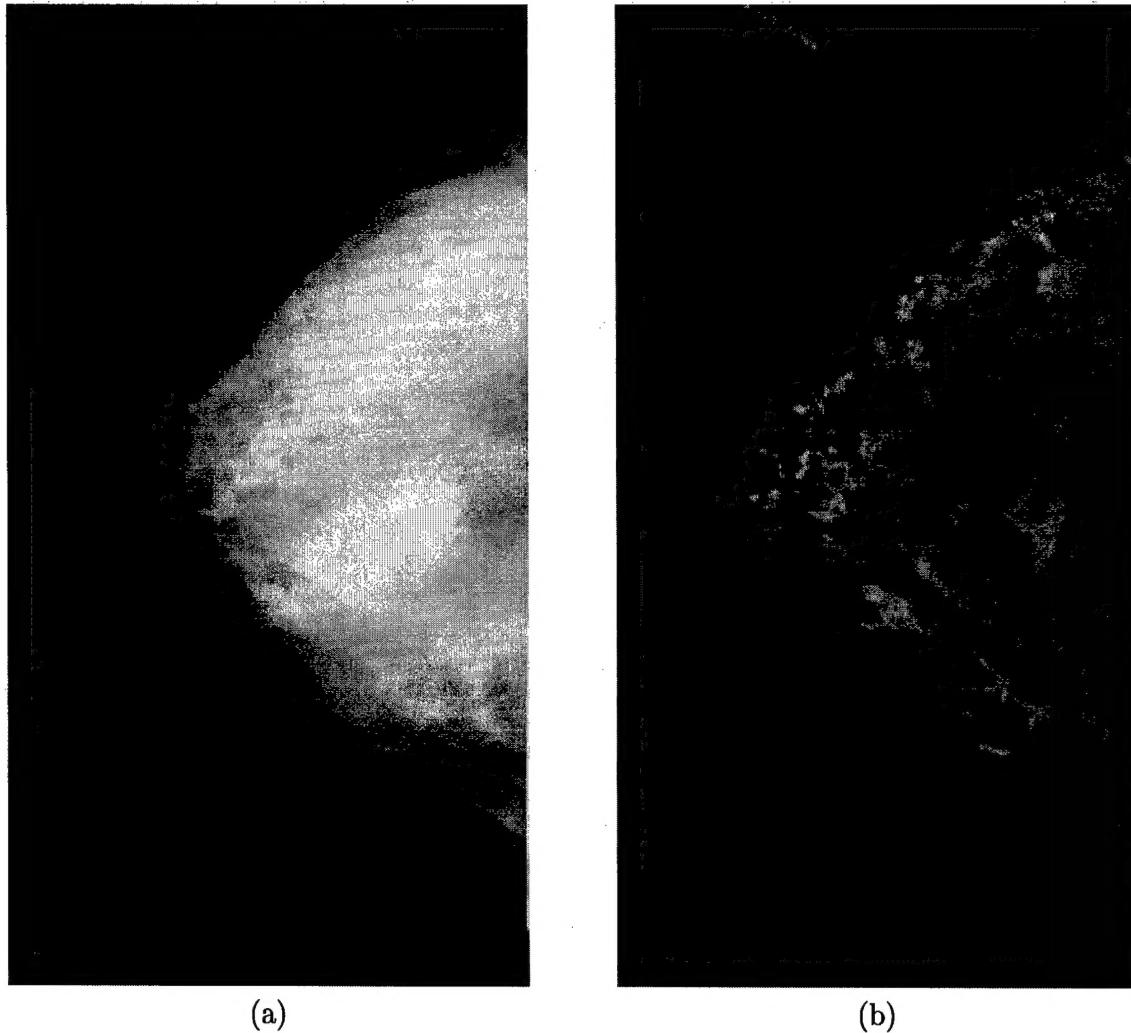


Figure 2. (a) The cranio-caudal view of the left breast. (b) Enhanced image improves visualization of the borders of the mass.

4. FUSION OF ENHANCED FEATURES

After coefficients are processed for enhancement of distinct mammographic features, the corresponding coefficients are combined according to a fusion rule into a new set of transform coefficients from which the fused result is reconstructed. As a fusion rule, the maximum oriented energy criterion was chosen: at each position and scale of the transform, the coefficient with greatest local energy was selected [12].

It is also possible to put distinct weights on selected features, and exclude other features from the final result.

Figure 2 shows the original mammogram and the processed image with improved contrast between the fat and glandular tissue.

5. CONCLUSION

The presented method incorporates a variety of properties of mammographic image enhancement techniques tailored to specific signs of malignancy into a unified computational framework. A multiscale spline derivative-based transform proved flexible enough for implicit enhancement of individual types of mammographic features and thus enabled processing within a single wavelet transform decomposition. In addition to its efficiency, the algorithm is well suited for further refinements; optimizations can be performed for each type of malignancy alone, and separately for the fusion strategy.

Our preliminary experiments imply that an enhancement via fusion approach can provide more obvious clues for radiologists. Further clinical tests are planned to verify that the versatility of this paradigm can provide a better viewing environment for a more reliable interpretation in screening mammography.

REFERENCES

- [1] A. F. Laine, S. Schuler, J. Fan, and W. Huda, "Mammographic feature enhancement by multiscale analysis," *IEEE Trans. Med. Imaging*, vol. 13, pp. 725-740, 1994.
- [2] J. Fan and A. Laine, "Multiscale contrast enhancement and denoising in digital radiographs," in *Wavelets in Medicine and Biology*, A. Aldroubi and M. Unser, Eds., CRC Press, Boca Raton, FL, 1996, pp. 163-189.
- [3] C.-M. Chang and A. Laine, "Enhancement of mammograms from oriented information," in *Proc. IEEE Int. Conf. Image Process.*, Santa Barbara, CA, Oct. 1997, vol. 3, pp. 524-527.
- [4] W. M. Morrow, R. B. Paranjape, R. M. Rangayyan, and J. E. L. Desautels, "Region-based contrast enhancement of mammograms," *IEEE Trans. Med. Imaging*, vol. 11, pp. 392-406, 1992.
- [5] R. N. Strickland and H. I. Hahn, "Wavelet transforms for detecting microcalcifications in mammograms," *IEEE Trans. Med. Imaging*, vol. 15, pp. 218-229, 1996.
- [6] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 891-906, 1991.
- [7] I. Koren, A. Laine, and F. Taylor, "Enhancement via fusion of mammographic features," in *Proc. IEEE Int. Conf. Image Process.*, Chicago, IL, Oct. 1998, vol. 1.
- [8] L. W. Estevez and N. D. Kehtarnavaz, "Computer assisted enhancement of mammograms for detection of microcalcification," in *Proc. IEEE Symp. Comput. Based Med. Syst.*, 1995, pp. 16-23.
- [9] H. Li, K. J. R. Liu, and S. C. B. Lo, "Fractal modeling of mammogram and enhancement of microcalcifications," *Proc. IEEE Nucl. Sci. Symp. Med. Imaging Conf.*, 1997, vol. 3, pp. 1850-1854.
- [10] M. Unser, A. Aldroubi, and S. J. Schiff, "Fast implementation of the continuous wavelet transform with integer scales," *IEEE Trans. Signal Process.*, vol. 42, pp. 3519-3523, 1994.
- [11] W. P. Kegelmeyer, J. M. Pruneda, P. D. Bourland, A. Hillis, M. V. Riggs, and M. L. Nipper, "Computer-aided mammographic screening for spiculated lesions," *Radiology*, vol. 191, pp. 331-337, 1994.
- [12] I. Koren, A. Laine, and F. Taylor, "Image fusion using steerable dyadic wavelet transform," in *Proc. IEEE Int. Conf. Image Process.*, Washington, D.C., Oct. 1995, vol. 3, pp. 232-235.